Fully Distributed Transactive Control Considering Pricing Dynamics and Network Constraints

Eduardo Mojica-Nava, *Member, IEEE*, Fredy Ruiz, *Senior Member, IEEE*, and Eder Baron-Prada, *Member, IEEE*

***Abstract*—Distributed energy resources have become a key element towards a smarter grid. However, several significant challenges for real-time implementation have emerged, in particular control design and its integration with distribution systems. In this work, a fully distributed dynamic transactive control to coordinate distributed energy resources in a distribution system considering physical network constraints based on saddle-point dynamics with a predictive-sensitivity conditioning term is proposed. An alternative interconnection of dynamical systems in different timescale preserving stability and optimality is introduced. A stability result for saddle-point dynamics with predictive-sensitivity conditioning is presented. Finally, simulation results for a distribution network with distributed energy resources adopted from an IEEE 37-bus test feeder are implemented to validate numerically the proposed approach.**

***Index Terms*—Transactive control, network constraints, distributed optimization.**

I. Introduction

Distributed energy resources (DERs) have become a key element towards a smarter grid. However, several significant challenges for real-time implementation of control algorithms and its integration with distribution networks have emerged. Traditional market-based techniques send pricing signals to the end-user to manage the demand response, such as real-time pricing [1] or time of use [2] approaches. On the other hand, a new market-based control technique, the so-called transactive energy systems (TES) [3], [4], [5], to integrate DERs through a market interaction while providing a scalable privacy-preserving framework has been gaining a great deal of attention for the coordination of numerous devices using value information under physical constraints [6], [7]. In [7], it is presented an important review of the design and analysis based on this framework. Recently in [8], a review of TES for microgrids is presented, and in [9] the TES framework is used for the optimal energy management of a multi-microgrid system. Traditional optimization methods have been proposed to model and solve the TES control problem [4], [5], [10], [11], [12], mostly at the transmission system level [4], [10], [13]. However, centralized solution in systems with many devices often presents computational and communication issues.

In the context of non-centralized solutions in TES, peer-to-peer (P2P) energy trading concept is gaining attention as a promising alternative [14], [15]. P2P energy trading can be implemented in a centralized market with a central network manager or decentralized based on agent's interactions. P2P-based methods allow DERs to trade energy in local energy markets with their neighbors. However, in general a full connected communication network is needed and the physical network constraints and network losses are not considered. Several P2P approaches have been discussed, from auction-based energy sharing models for distribution systems to the recent dynamic energy pricing dynamics strategies based on game-theoretic two-stage models. In a two-stage approach, DERs determine the amount of energy to share/consume in the first stage and the energy price is found in the second stage. A social welfare maximization is solved to set the market clearing price (MCP) [16]. Recently in [15], a P2P transactive market considering physical network constraints is proposed based on a distributed alternate direction method of multipliers with a central network manager. Network constraints such as power flow and bus voltage [17], [18] has not been traditionally included in energy trading in distribution networks. In the few works where the network constraints are considered, a centralized mechanism has been used.

In response to the centralized and hierarchical limitations, there has been an increasing interest in applying distributed optimization strategies based on the consensus on their estimates using local information in discrete-time. In spite of the recent success of such algorithms, several limitations arise in real-time operation: sensitivity to environmental variations, reliance on synchronized information between agents, or a two-time scale allocation problem [19]. In contrast, in continuous-time optimization, saddle-point dynamics have become an important strategy in the design of distributed optimal feedback controllers [20]. Since the seminal

Manuscript received 14 September 2021; revised 22 February 2022, 16 May 2022, and 20 June 2022; accepted 21 June 2022. Date of publication 24 June 2022; date of current version 20 February 2023. Paper no. TSG-01489-2021. *(Corresponding author: Eduardo Mojica-Nava.)*

Eduardo Mojica-Nava is with the Department of Electrical and Electronics Engineering, Universidad Nacional de Colombia, Bogotá 111321, Colombia, and also with the Department of Electronics, Informatics, and Bioengineering, Politecnico de Milano, 20133 Milan, Italy (e-mail: eamojican@unal.edu.co).

Fredy Ruiz is with the Department of Electronics, Informatics, and Bioengineering, Politecnico de Milano, 20133 Milan, Italy.

Eder Baron-Prada was with the Computer, Electrical and Mathematical Sciences and Engineering Division (CEMSE), King Abdullah University of Science and Technology, Thuwal 23955, Saudi Arabia. He is now with the Center for Energy, Austrian Institute of Technology, 1210 Vienna, Austria, and also with the Automatic Control Laboratory, ETH Zurich, 8092 Zürich, Switzerland (e-mail: ebaron@ethz.ch).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSG.2022.3186163.

Digital Object Identifier 10.1109/TSG.2022.3186163

works [21] [22], the use of saddle-point dynamics, also known as primal-dual dynamics, to obtain the saddle-points of the Lagrangian in constrained optimization has been a fundamental algorithm. Nevertheless, the renovated emerging attention on saddle-point dynamics is mainly motivated for network optimization problems where the objective function is an aggregation of local functions with local constraints, then the saddle-point dynamics for such problems is essentially compliant to distributed implementation [23], [24].

On the other, it has been shown that the TES framework can be modeled as a bi-level optimization problem, where the upper level is the determination stage where the production and demand level of energy of each transactive agent is obtained. The lower level is the pricing mechanism, usually solved by a network manager or a distribution system operator. An equilibrium problem of bi-level optimization resulting in equilibrium problems with equality constraints (EPEC) are in general intractable or difficult to solve due to its nested nature. To deal with EPEC computational issues in [25] is proposed a single-level electricity market equilibrium model, where three different single-level optimization problems are proposed (consumer, producer, and market operator) and solved jointly using concepts of complementarity [26] in a centralized fashion. Considering the exposed limitations and challenges in the implementation of DERs in the distribution systems, the main contribution of this work is presented in the next section.

A. Statement of Contributions

The main contribution of this work is a fully distributed dynamic transactive control to coordinate DERs in a distribution system considering physical network constraints. We propose a fully distributed continuous-time optimization scheme, interconnecting two levels of decision based on two different timescales into a dynamical system in a single timescale. This alternative interconnection keeps a single timescale and preserves stability and optimality. Contrary to traditional two timescales separation algorithms, the proposed dynamics overcome the limitation of the slow convergence rate due to the nested nature and its computational issues. In addition, the physical network constraints are included in the distribution system model. Fully distributed computations imply that there is no a central coordinator, and instead the DERs are interacting through a not fully connected communication network in real-time, responding to the computational, cyber-secure, and communication issues of centralized or hierarchical solutions. Finally, two simulation cases for a distribution network with distributed energy resources adopted from an IEEE 37-bus test feeder are implemented to validate numerically the proposed approach.

The proposed distributed interconnected system and its fundamental features are summarized as follows. In the first decision level, the faster-time scale network, a distributed optimization algorithm in continuous-time is proposed based on a saddle-point dynamics [20] for solving a power allocation optimization problem between DERs using the pricing information. In the second decision level, the slower timescale network, agents solve also a welfare optimization problem by

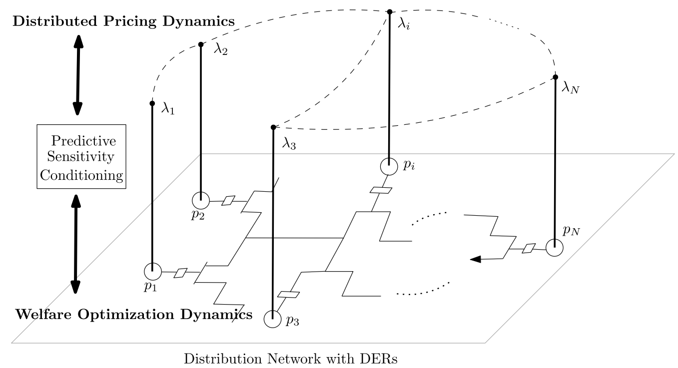

Fig. 1. Distributed transactive energy system.

means of a distributed pricing dynamics using a saddle-point dynamics, considering as a constraint the optimal solution of the faster-level and the physical network constraints to obtain the optimal energy prices for a market period. The inclusion of the physical network constraints such as the power flow and the bus voltage in a fully distributed energy trading is an advantage in real-time integration of DERs in distribution networks. Hence, we have a two-level dynamic transactive control involving a feedback loop system solving a bi-level optimization problem in continuous-time. In order to integrate the dynamical systems in two timescales in a single timescale dynamical system, we use a predictive-sensitivity conditioning term. The proposed predictive-sensitivity conditioning dynamics can be interpreted as a feed-forward term in the faster system anticipating the dynamics of the slower dynamics using a sensitivity conditioning matrix [27]. A general scheme of the proposed distributed transactive energy system is presented in Fig. 1. Finally, the proposed dynamics are supposed to consider optimization algorithms as feedback dynamical systems, which opens up new possibilities to bestow complex real-world systems in a robust feedback-based optimization.

The rest of the paper is organized as follows. In Section II, a bi-level optimization problem is described to represent the distribution system with DERs and network constraints in a transactive energy framework; a DER power allocation problem is presented with the definition of the utility functions. Section III describes a fully distributed feedback-based optimization transactive control with predictive-sensitivity conditioning transactive control algorithm based on saddle-point dynamics. In Section IV, stability and convergence results are presented. In Section V, simulation results are presented to illustrate the effectiveness of the proposed algorithm. Finally, in Section V some conclusions and future work are drawn.

B. Preliminaries and Notation

A communication graph is denoted as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. DERs are represented as a set of nodes $\mathcal{V} = \{1, 2, \dots, V\}$. $\mathcal{E}$ is the set of links between agents, i.e., $(i, j) \in \mathcal{E}$ if there is a link between $i \in \mathcal{V}$ and $j \in \mathcal{V}$. $\mathcal{N}_i$ is the neighborhood of i-th agent. If there is at least one path between two different nodes, the graph $\mathcal{G}$ is connected. $L_{\mathcal{G}}$ is the Laplacian of the communication graph between DERs.

II. Problem Statement

In this section, a DER power allocation problem with the utility functions for each type of DER (generator or consumer) is presented. A bidirectional connected graph is assumed as the information structure between DERs. The dynamic transactive control is a feedback loop system solving a bi-level optimization problem, with two levels of decision based on two different time-scales. In the first level, the faster-time scale network, a power allocation optimization problem is stated between DERs using the pricing information. In the slower timescale network, agents solve also a welfare optimization problem, considering as a constraint the optimal solution of the faster-level and the physical network constraints to obtain the optimal energy prices for a market period.

A. DER Power Allocation Problem

The DERs consist of a set of adjustable loads (or smart consumer) and a set of distributed generators. First, we introduce smart consumers. Let k be an adjustable load such as $k = \{1, 2, \dots, K\}$, with payoff function defined as

$$U_k^d(\lambda_k, p_k) = V_k(p_k) - \lambda_k p_k, \tag{1}$$

where V_k are the valuation function for each consumer representing the monetary value that this consumer assigns to the consumption of energy, and λ_k is the energy price for consumer k [1]. The adjustable loads have a consumption limit as $\underline{p}_k \leq p_k \leq \overline{p}_k$. It should be noticed that the minimum adjustable load is the base load needed for the consumer. Some conditions should be imposed on the valuation function such as it has to be twice differentiable, concave, non-decreasing, and $V_k(0) = 0$. For instance, consider a quadratic valuation function as $V_k(p_k) = \alpha_k p_k^2 + \beta_k p_k$, where α_k and β_k are utility coefficients belonging to $\mathbb{R}_+$.

Similar to the adjustable loads, consider a set of distributed generators $G = \{1, 2, \dots, M\}$, each distributed generator has a payoff function:

$$U_j^g\big(\lambda_j, p_j\big) = \lambda_j p_j - C_j\big(p_j\big) \tag{2}$$

where p_j is the amount of power produced by each generator j, λ_j is the energy price for the generator j, and $C_j(p_j)$ is the cost function for producing energy for each generator. The cost function is traditionally chosen as a convex quadratic function such as $C_j = \alpha_j p_j^2 + \beta_j p_j$, where α_j, β_j are positive cost coefficients. The distributed generators have a minimum and maximum generation limit $\underline{p}_j \leq p_j \leq \overline{p}_j$.

The power of each DER is stacked in the vector $\boldsymbol{p} = [p_1, \dots, p_N]^\top$, where the total number of agents is $N = K+M$, and the set of interacting agents is $i \in D = \{1, 2, \dots, N\}$.

The previous described agents are located in a power distribution network with a set of buses including the distributed energy resources $m \in I = \{0, 1, \dots, I\}$, i.e., generators and consumers.

These buses are interconnected between them, inducing a set of line segments $S = \{(m, l) | m \in I, l \in I\} \subseteq I \times I$. It is assumed the bus 0 as the reference bus with 1 p.u. voltage. v_m is the voltage of bus m. Let define r_{ml} and x_{ml} as the resistance and reactance of the line (m, l), respectively, and p_{ml} as the real power flow through this line. The neighbors of bus m are denoted as $\mathcal{N}_m$. Assuming some standard conditions as in [15] and a constant reactive power, we can obtain a linearized DisFlow model as follows,

$$\boldsymbol{p} = W^\top \mathbf{P} \tag{3}$$

where $\boldsymbol{p}$ is the vector of production/consumption of real power. $\mathbf{P} = \{p_{12} \;\; p_{13} \;\; \cdots p_{ml} \;\; \cdots\}$ is the vector of power flows p_{ml}. The matrix W is the graph incidence matrix with entries $[W]_{lm} = 1$ if a line l is leaving the bus m and $[W]_{lm} = -1$ if a line l enters the bus m. The bus voltage vector is as follows

$$\mathbf{v} = R\boldsymbol{p} + \overline{\mathbf{v}} \tag{4}$$

with $\overline{\mathbf{v}} = v_0\mathbf{1} + X\mathbf{q}$, where $\mathbf{q}$ is a fixed reactive power vector, and R and X are resistance and reactance matrices, respectively.

Considering (3) and (4), and stacked the network constraint variables $\mathbf{v}$ and $\mathbf{P}$ as $\boldsymbol{\theta} = [\mathbf{v} \; \mathbf{P}]^\top$, it is possible to obtain in compact form $\boldsymbol{h}(\boldsymbol{p}, \boldsymbol{\theta})$ as

$$\boldsymbol{h}(\boldsymbol{p}, \boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{p} - R^{-1}(\boldsymbol{v} - \overline{\mathbf{v}}) \\ \boldsymbol{p} - W^\top \mathbf{P} \end{bmatrix}. \tag{5}$$

Furthermore, the upper and lower limits of the active power $\underline{p}_i \leq p_i \leq \overline{p}_i$, $\forall\; i \in D$, voltage at each node $\underline{v}_m \leq v_m \leq \overline{v}_m$, $\forall\; m \in I$, and power flows $\underline{p}_{ml} \leq p_{ml} \leq \overline{p}_{ml}$, $\forall\; (m, l) \in I$ can be arranged in a vector of inequality constraints $\boldsymbol{g}(\boldsymbol{p}, \boldsymbol{\theta}) \leq \mathbf{0}$.

For a given fixed energy price $\bar{\lambda}$, the DERs power allocation problem is formulated to maximize the benefits of DERs under constraints as follows

$$\begin{aligned} \max_{\boldsymbol{p}, \boldsymbol{\theta}} \quad & U\big(\bar{\lambda}, \boldsymbol{p}\big) = \sum_{k=1}^{K} U_k^d\big(p_k, \bar{\lambda}\big) + \sum_{j=1}^{M} U_j^g\big(p_j, \bar{\lambda}\big) \\ \text{s.t.} \quad & \boldsymbol{h}(\boldsymbol{p}, \boldsymbol{\theta}) = \mathbf{0} \\ & \boldsymbol{g}(\boldsymbol{p}, \boldsymbol{\theta}) \leq \mathbf{0}. \end{aligned} \tag{6}$$

B. Pricing Optimization Problem

The main objective of the pricing problem is to determine the market clearing price (MCP) between energy producer and smart consumers. Traditionally, the distribution system operator (DSO) solves a static optimization problem that maximize the welfare of the population over a market cycle. It is noticed that the coordination architecture is a Stackelberg game, which instead of simultaneous play, it exhibits a hierarchical order of play conforming a hierarchical decision-making structure, which is a challenging bi-level optimization problem. In this case, the leader is the DSO and the DERs in the distribution network are the followers.

In the objective function, it is assumed that the DERs have already chosen the optimal generation/consumption profile for a specific price λ, $p^*(\lambda)$. The optimal market clearing price λ^* is then obtained by solving the following bi-level optimization problem

$$\begin{aligned} \max_{\lambda} \quad & U\big(\lambda, \boldsymbol{p}^*\big) \\ \text{s.t.} \quad & \boldsymbol{p}^* = \arg\max_{\boldsymbol{p}, \boldsymbol{\theta}} \{U(\lambda, \boldsymbol{p}) : \boldsymbol{h}(\boldsymbol{p}, \boldsymbol{\theta}) = \mathbf{0}, \boldsymbol{g}(\boldsymbol{p}, \boldsymbol{\theta}) \leq \mathbf{0}\}, \end{aligned} \tag{7}$$

where $\boldsymbol{h}(\boldsymbol{p},\boldsymbol{\theta})=\boldsymbol{0}$ are the compact matrix form of the physical network constraints in Problem (6) defined in (5). As a result, the DSO has as constraint another optimization problem based on the payoff functions of DERs. It should be noticed that not only the structure of the optimization problem is hierarchical in order of play, but also the leader and the followers are in different time-scales.

III. Distributed Feedback-Based Optimization Transactive Control

We introduce a distributed transactive control algorithm considering distributed power integration, flexible power demand, physical network constraints, and power losses allocation using a feedback-based distributed optimization algorithm in continuous-time. A general scheme of the bi-level distributed transactive energy system is presented in Fig. 1. In order to solve the problem without a leader agent, i.e., without DSO, we propose a two-level distributed problem where DERs have to obtain the MCP first (slower problem) and then adjust the consumption/generation profiles (faster problem). As it has been mentioned, there is a hierarchical decision-making structure, where DERs exchange information through a communication network. Some assumptions have to be considered to assure optimality and existence of solutions of Problem (7). First, the feasibility of the initial solution have to be assumed to guarantee the stability of the dynamics.

Assumption 1: The decision variables start inside the feasible set.

The connectivity of the communication graph is needed to guarantee the information exchange between all DERs.

Assumption 2: The communication graph between DERs is connected.

Finally, it is needed the regularity of the welfare functions for numerical stability of the proposed dynamics.

Assumption 3: Every payoff function $U_k^d(\lambda_k,p_k)$ and $U_j^g(\lambda_j,p_j)$ is assumed to be differentiable everywhere and with Lipschitz continuous partial derivatives, and the Hessian $\nabla^2_{\boldsymbol{pp}}U(\lambda,\boldsymbol{p})$ is globally invertible.

We propose a continuous-time distributed optimization scheme with two levels of decision based on two different timescales. For uniformity in stating the Karush–Kuhn–Tucker (KKT) conditions, we use minimization instead of maximization by changing the payoff function as,

$$J(\boldsymbol{\lambda},\boldsymbol{p})=-U(\boldsymbol{\lambda},\boldsymbol{p})$$

where we use the stacked vector of prices for each DER as $\boldsymbol{\lambda}=[\lambda_1,\ldots,\lambda_N]^\top$. In the slower timescale network, agents solve also an optimization problem (7) considering as a constraint the optimal solution of the faster level with physical network constraints, and communication network between agents to guarantee the price consensus to obtain the optimal energy price for a market period as a distributed pricing optimization problem:

$$\begin{aligned}\min_{\boldsymbol{\lambda}}\ & J\big(\boldsymbol{\lambda},\boldsymbol{p}^*\big)\\ \text{s.t.}\ & \boldsymbol{p}^*=\arg\min_{\boldsymbol{p}}\{J(\boldsymbol{\lambda},\boldsymbol{p}):\boldsymbol{h}(\boldsymbol{p},\boldsymbol{\theta})=\boldsymbol{0},\boldsymbol{g}(\boldsymbol{p},\boldsymbol{\theta})\le\boldsymbol{0}\},\\ & L_{\mathcal{G}}\boldsymbol{\lambda}=\boldsymbol{0},\end{aligned}\tag{8}$$

where $L_{\mathcal{G}}$ is the Laplacian of the communication graph between DERs. The corresponding Lagrangian function for the upper-level optimization problem is,

$$L_1\big(\boldsymbol{\lambda},\boldsymbol{p}^*(\boldsymbol{\lambda}),\boldsymbol{\gamma}\big)=J\big(\boldsymbol{\lambda},\boldsymbol{p}^*(\boldsymbol{\lambda})\big)+\boldsymbol{\gamma}^\top L_{\mathcal{G}}\boldsymbol{\lambda},\tag{9}$$

where $\boldsymbol{\gamma}$ is the vector of Lagrange multipliers associated to the consensus constraint. The corresponding Lagrangian function for a lower-level optimization problem is defined by,

$$L_2\big(\boldsymbol{\lambda}^*,\boldsymbol{p},\boldsymbol{\theta},\boldsymbol{\mu}\big)=J\big(\boldsymbol{\lambda}^*,\boldsymbol{p}\big)+\boldsymbol{\mu}^\top\boldsymbol{h}(\boldsymbol{p},\boldsymbol{\theta})+\boldsymbol{\tau}^\top\log(-\boldsymbol{g}(\boldsymbol{p},\boldsymbol{\theta})),\tag{10}$$

where $\boldsymbol{\mu}$ is the vector of Lagrange multipliers. We adopt the log barrier, as in several interior point methods, whereby using the log function we impose a barrier that increases if the variable is in a close neighborhood. The sharpness of the barrier can be tuned by the $\boldsymbol{\tau}$ parameter that is tuned in order to have a stronger or relaxed condition over the inequality constraints.

Local solution concepts and first and second order necessary and sufficient conditions for bi-level optimization problems [28] are needed to establish the saddle-point dynamics with predictive-sensitivity conditioning.

Definition 1: A point $(\boldsymbol{\lambda}^*,\boldsymbol{\gamma}^*,\boldsymbol{p}^*,\boldsymbol{\theta}^*,\boldsymbol{\mu}^*)$ is said to be a local solution to the bi-level optimization problem (8) if

(a) $\boldsymbol{p}^*$ is a local minimum of $L_2(\boldsymbol{\lambda}^*,\boldsymbol{\gamma}^*,\boldsymbol{p}^*,\boldsymbol{\theta}^*,\boldsymbol{\mu}^*)$;

(b) there exists a neighborhood $\Omega\in\mathbb{R}^N$ of $(\boldsymbol{\lambda}^*,\boldsymbol{\gamma}^*,\boldsymbol{p}^*,\boldsymbol{\theta}^*,\boldsymbol{\mu}^*)$ such that $J(\boldsymbol{\lambda}^*,\boldsymbol{\gamma}^*,\boldsymbol{p}^*,\boldsymbol{\theta}^*,\boldsymbol{\mu}^*)\le J(\boldsymbol{\lambda}^*,\boldsymbol{\gamma}^*,\boldsymbol{p},\boldsymbol{\theta}^*,\boldsymbol{\mu}^*)$ for all local solutions $(\boldsymbol{\lambda}^*,\boldsymbol{\gamma}^*,\boldsymbol{p}^*,\boldsymbol{\theta}^*,\boldsymbol{\mu}^*)\in\Omega$

Note that $(\boldsymbol{\lambda}^*,\boldsymbol{\gamma}^*,\boldsymbol{p}^*,\boldsymbol{\theta}^*,\boldsymbol{\mu}^*)$ is said to be a solution if Ω is all the subspace is defined by the box constraints.

Lemma 1 (First-Order Optimality Conditions): If $(\boldsymbol{\lambda}^*,\boldsymbol{p}^*,\boldsymbol{\theta}^*)$ is a local solution of Problem (8) with Lagrange multipliers $\boldsymbol{\mu}^*$ and $\boldsymbol{\gamma}^*$, then it is a stationary point satisfying the first-order KKT conditions

$$\begin{aligned}\nabla_{\boldsymbol{\lambda}}L_1\big(\boldsymbol{\lambda}^*,\boldsymbol{\gamma}^*,\boldsymbol{p}^*(\boldsymbol{\lambda}^*)\big)&=0,\\ \nabla_{\boldsymbol{\gamma}}L_1\big(\boldsymbol{\lambda}^*,\boldsymbol{\gamma}^*,\boldsymbol{p}^*(\boldsymbol{\lambda}^*)\big)&=0,\\ \nabla_{\boldsymbol{p}}L_2\big(\boldsymbol{\lambda}^*,\boldsymbol{\gamma}^*,\boldsymbol{p}^*,\boldsymbol{\theta}^*,\boldsymbol{\mu}^*\big)&=0,\\ \nabla_{\boldsymbol{\theta}}L_2\big(\boldsymbol{\lambda}^*,\boldsymbol{\gamma}^*,\boldsymbol{p}^*,\boldsymbol{\theta}^*,\boldsymbol{\mu}^*\big)&=0,\\ \nabla_{\boldsymbol{\mu}}L_2\big(\boldsymbol{\lambda}^*,\boldsymbol{\gamma}^*,\boldsymbol{p}^*,\boldsymbol{\theta}^*,\boldsymbol{\mu}^*\big)&=0.\end{aligned}\tag{11}$$

Lemma 2 (Second-Order Optimality Conditions): If a stationary point $(\boldsymbol{\lambda}^*,\boldsymbol{p}^*,\boldsymbol{\theta}^*)$ with Lagrange multipliers $\boldsymbol{\mu}^*$ and $\boldsymbol{\gamma}^*$ and satisfies

$$\begin{aligned}\nabla^2_{\boldsymbol{\lambda\lambda}}L_1\big(\boldsymbol{\lambda}^*,\boldsymbol{\gamma}^*,\boldsymbol{p}^*(\boldsymbol{\lambda}^*)\big)&\ge 0,\\ \nabla^2_{\boldsymbol{\gamma\gamma}}L_1\big(\boldsymbol{\lambda}^*,\boldsymbol{\gamma}^*,\boldsymbol{p}^*(\boldsymbol{\lambda}^*)\big)&\le 0,\\ \nabla^2_{\boldsymbol{pp}}L_2\big(\boldsymbol{\lambda}^*,\boldsymbol{\gamma}^*,\boldsymbol{p}^*,\boldsymbol{\theta}^*,\boldsymbol{\mu}^*\big)&\ge 0,\\ \nabla^2_{\boldsymbol{\theta\theta}}L_2\big(\boldsymbol{\lambda}^*,\boldsymbol{\gamma}^*,\boldsymbol{p}^*,\boldsymbol{\theta}^*,\boldsymbol{\mu}^*\big)&\ge 0,\\ \nabla^2_{\boldsymbol{\mu\mu}}L_2\big(\boldsymbol{\lambda}^*,\boldsymbol{\gamma}^*,\boldsymbol{p}^*,\boldsymbol{\theta}^*,\boldsymbol{\mu}^*\big)&\le 0,\end{aligned}\tag{12}$$

Then, $(\boldsymbol{\lambda}^*,\boldsymbol{p}^*,\boldsymbol{\theta}^*)$ is a local solution of the bi-level optimization Problem (8).

We stack the variable as $z=[\boldsymbol{p}\quad\boldsymbol{\theta}\quad\boldsymbol{\mu}]^\top$ for the faster subsystem (lower-level). Assumption 3 guarantees that, for a given $\boldsymbol{\lambda}$, the lower-level problem in (8) has at most a single optimal solution z^* satisfying the first-order optimality conditions of the lower-level $\nabla_z L_2(\boldsymbol{\lambda}, z_2^*) = 0$. The implicit function theorem [29] guarantees the local existence of the optimal solution function $z^*(\boldsymbol{\lambda})$, and its derivative is obtained as

$$\nabla_{\boldsymbol{\lambda}} z^*(\boldsymbol{\lambda}) = -\left(\nabla^2_{zz} L_2\big(\boldsymbol{\lambda}, z^*(\boldsymbol{\lambda})\big)\right)^{-1} \nabla^2_{z\boldsymbol{\lambda}} L_2\big(\boldsymbol{\lambda}, z^*(\boldsymbol{\lambda})\big).$$

In addition, this optimal solution function can be used to give an expression for the directional derivative in points where $z = z^*(\boldsymbol{\lambda})$ and defined as,

$$\begin{aligned}\mathcal{D}_{\boldsymbol{\lambda}} L_1\big(\boldsymbol{\lambda}, \boldsymbol{\gamma}, z^*(\boldsymbol{\lambda})\big) &:= \nabla_{\boldsymbol{\lambda}} L_1\big(\boldsymbol{\lambda}, \boldsymbol{\gamma}, z^*(\boldsymbol{\lambda})\big) \\ &\quad + \nabla_{\boldsymbol{\lambda}} z^*(\boldsymbol{\lambda})^\top \nabla_z L_2\big(\boldsymbol{\lambda}, z^*(\boldsymbol{\lambda})\big). \end{aligned} \tag{13}$$

Considering Assumption 3, the analytic expression of (13) is well-defined at any point $(\boldsymbol{\lambda}, z)$ so that we can define an extended sensitivity for any point $(\boldsymbol{\lambda}, z)$ as follows

$$S^{\boldsymbol{p}}_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}, z) := -\left(\nabla^2_{zz} L_2(\boldsymbol{\lambda}, z)\right)^{-1} \nabla^2_{z\boldsymbol{\lambda}} L_2(\boldsymbol{\lambda}, z), \tag{14}$$

while satisfying

$$S^{\boldsymbol{p}}_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}, z)|_{(\boldsymbol{\lambda}, z^*(\boldsymbol{\lambda}))} = \nabla_{\boldsymbol{\lambda}} z^*(\boldsymbol{\lambda}).$$

The directional derivative using the sensitivity matrix is extended to any point $(\boldsymbol{\lambda}, z)$ as

$$\begin{aligned}\mathcal{D}_{\boldsymbol{\lambda}} L_1(\boldsymbol{\lambda}, \boldsymbol{\gamma}, z) &:= \nabla_{\boldsymbol{\lambda}} L_1(\boldsymbol{\lambda}, \boldsymbol{\gamma}, z) \\ &\quad + S^{\boldsymbol{p}}_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}, z)^\top \nabla_z L_2(\boldsymbol{\lambda}, z). \end{aligned} \tag{15}$$

In the next section, it is presented the main contribution of this work, i.e., the saddle-point dynamics including pricing and power allocation in a single timescale using the predictive-sensitivity conditioning matrix.

A. Predictive-Sensitivity Conditioning Dynamics

In this section, we present the predictive-sensitivity conditioning technique and saddle-point dynamics to solve the bi-level optimization problem (8). The saddle-point dynamics, gradient descent in some variables and gradient ascent in the others, are obtained following traditional ideas of gradient dynamics for convex optimization problems with equality constraints [30].

Considering the Lagrangian saddle function of the upper-level (9), we obtain a distributed pricing dynamics (see the upper-level in Fig. 1) that seek to minimize the price $\boldsymbol{\lambda}$ (gradient descent for consumers and gradient ascent for generators) using saddle-point dynamics as follows

$$\Sigma_1 : \begin{cases} \dot{\boldsymbol{\lambda}} = -\mathcal{D}_{\boldsymbol{\lambda}} L_1(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \boldsymbol{p}), \\ \dot{\boldsymbol{\gamma}} = L_{\mathcal{G}} \boldsymbol{\lambda}. \end{cases} \tag{16}$$

On the other hand, in the faster timescale network, an optimization algorithm in continuous-time is proposed based on saddle-point dynamics for solving the optimization problem (8) between DERs using the price information from the upper-level (see the lower-level in Fig. 1). In this case, we seek to minimize in the power variable $\boldsymbol{p}$ considering the network constraints $\boldsymbol{h}(\boldsymbol{p}, \boldsymbol{\theta}) = 0$ to obtain a saddle-point

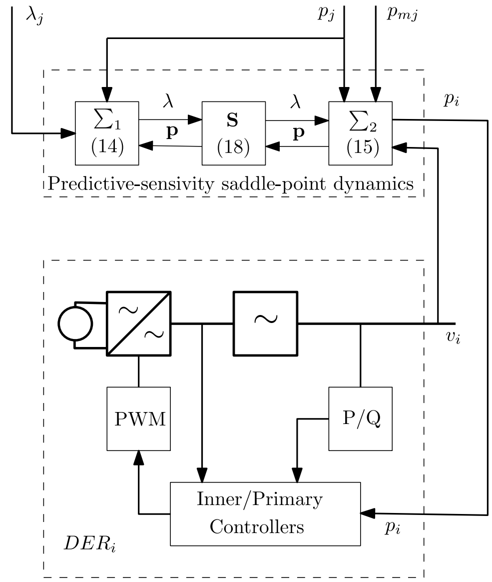

Fig. 2. Block diagram of the predictive-sensivity saddle-point dynamics (23) interconnected to DER_i, where $(p_j, \lambda_j) \in \mathcal{N}_i$ indicates signals from DERs in the neighborhood.

dynamics as follows

$$\Sigma_2 : \begin{cases} \dot{\boldsymbol{p}} = -\nabla_{\boldsymbol{p}} L_2(\boldsymbol{\lambda}, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}) \\ = -\nabla_{\boldsymbol{p}} J(\boldsymbol{\lambda}, \boldsymbol{p}) + \boldsymbol{\mu} + \tau \nabla_{\boldsymbol{p}} \log(-\boldsymbol{g}(\boldsymbol{p}, \boldsymbol{\theta})), \\ \dot{\boldsymbol{\theta}} = -\nabla_{\boldsymbol{\theta}} L_2(\boldsymbol{\lambda}, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}) \\ = -\boldsymbol{\mu}^\top \nabla_{\boldsymbol{\theta}} \boldsymbol{h}(\boldsymbol{p}, \boldsymbol{\theta}) - \tau \nabla_{\boldsymbol{\theta}} \log(-\boldsymbol{g}(\boldsymbol{p}, \boldsymbol{\theta})), \\ \dot{\boldsymbol{\mu}} = \nabla_{\boldsymbol{\mu}} L_2(\boldsymbol{\lambda}, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}) = \boldsymbol{h}(\boldsymbol{p}, \boldsymbol{\theta}). \end{cases} \tag{17}$$

In order to solve the two-level interconnected dynamical system represented by Σ_1 and Σ_2 in a single timescale, avoiding the inherent complications of the traditional nested optimization, we introduce the predictive-sensitivity conditioning method. The predictive-sensitivity conditioning adds a feed-forward term based on the optimization sensitivity to the saddle-point dynamics of the faster system Σ_2 (17) to predict the dynamics of the slower saddle-point dynamics Σ_1 (16). To present the theoretical results on predictive-sensitivity conditioning in compact form, we use the stacked variables of each timescale as (x, z), where $x = [\boldsymbol{\lambda} \quad \boldsymbol{\gamma}]^\top$. The gradient flows are represented by

$$G_1(x, z) := \begin{bmatrix} -\mathcal{D}_{\boldsymbol{\lambda}} L_1 \\ L_{\mathcal{G}} \boldsymbol{\lambda} \end{bmatrix}, \tag{18}$$

and

$$G_2(x, z) := \begin{bmatrix} -\nabla_{\boldsymbol{p}} L_2 \\ -\nabla_{\boldsymbol{\theta}} L_2 \\ \nabla_{\boldsymbol{\mu}} L_2 \end{bmatrix}, \tag{19}$$

respectively. Based on a recent work [31], we use a predictive-sensitivity conditioning interconnection between the two timescale systems without the need of the timescale separation. This interconnection preserves the optimal solution $(\boldsymbol{\lambda}^*, z^*(\boldsymbol{\lambda}^*)) = (\boldsymbol{\lambda}^*, \boldsymbol{p}^*(\boldsymbol{\lambda}^*), \boldsymbol{\theta}^*)$ and its convergence properties. A sensitivity based conditioning matrix $\boldsymbol{S}$ is used as the interconnection between both subsystems as shown in Fig. 2, where it is shown a DER with the control loop and the information received from the DERs in his neighborhood $\mathcal{N}_i$. The proposed feedback optimization control level, i.e., the predictive-sensitivity saddle-point dynamics, solve the MCP (λ^*) and the active power reference p_i^* dynamically, and these signals are sent to primary/inner controllers of each DER.

Definition 2: Consider the extended sensitivity $S_{\boldsymbol{\lambda}}^{\boldsymbol{p}}(\boldsymbol{\lambda}, z)$ defined in (14), the conditioning matrix is defined as

$$\boldsymbol{S} = \begin{bmatrix} I & 0 \\ -S_{\boldsymbol{\lambda}}^{\boldsymbol{p}}(\boldsymbol{\lambda}, z) & I \end{bmatrix} \tag{20}$$

where I is an identity matrix of the appropriate dimension.

Using Definition 2, we obtain the predictive-sensitivity conditioning system:

$$\boldsymbol{S}\begin{bmatrix} \dot{x} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} I & 0 \\ -S_{\boldsymbol{\lambda}}^{\boldsymbol{p}}(\boldsymbol{\lambda}, z) & I \end{bmatrix}\begin{bmatrix} \dot{x} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} G_1(x, z) \\ G_2(x, z) \end{bmatrix} \tag{21}$$

Notice that the extended sensitivity term modifies the dynamics of the faster system as

$$\dot{z} = G_2(x, z) + S_{\boldsymbol{\lambda}}^{\boldsymbol{p}}(\boldsymbol{\lambda}, z)G_1(x, z),$$

where G_2 drives z to the optimal solution, $z^*(\boldsymbol{\lambda})$, and the predictive-sensitivity feed-forward term $S_{\boldsymbol{\lambda}}^{\boldsymbol{p}}G_1$ anticipates the variation of $z^*(\boldsymbol{\lambda})$ due to the pricing dynamics $\dot{\boldsymbol{\lambda}}$.

In the original variables, the predictive-sensitivity conditioning $S_{\boldsymbol{\lambda}}^{\boldsymbol{p}}(\boldsymbol{\lambda}, z)$ (14) using the gradient of (18) and (19) is as follows

$$S_{\boldsymbol{\lambda}}^{\boldsymbol{p}} = -\left(\nabla_{\boldsymbol{pp}}^2 J(\boldsymbol{\lambda}, \boldsymbol{p})\right)^{-1}\left(\nabla_{\boldsymbol{p\lambda}}^2 J(\boldsymbol{\lambda}, \boldsymbol{p})\right) \tag{22}$$

Considering the payoff function $J(\boldsymbol{\lambda}, \boldsymbol{p})$, it is observed that the term $\nabla_{\boldsymbol{pp}}^2 J(\boldsymbol{\lambda}, \boldsymbol{p})$ is a diagonal matrix with factors of the coefficients α_i as its entries, and the term $\nabla_{\boldsymbol{p\lambda}}^2 J(\boldsymbol{\lambda}, \boldsymbol{p})$ is also a diagonal matrix with $+/-1$ depending on if it is a generator or a consumer as its entries so that the sensitivity function is a positive definite matrix of the form

$$S_{\boldsymbol{\lambda}}^{\boldsymbol{p}} = c\begin{bmatrix} 2\alpha_1 & 0 & \cdots & 0 \\ 0 & 2\alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 2\alpha_N \end{bmatrix},$$

where c is a constant depending on the coefficients. We have introduced all the fundamental ingredients to establish the main contribution of this work: a complete saddle-point dynamics in compact matrix form with the predictive-sensitivity conditioning term as follows

$$\begin{aligned} \dot{\boldsymbol{\lambda}} &= -\nabla_{\boldsymbol{\lambda}} J + S_{\boldsymbol{\lambda}}^{\boldsymbol{p}}\left(\nabla_{\boldsymbol{p}} L_2 + \nabla_{\boldsymbol{\theta}} L_2 - \nabla_{\boldsymbol{\mu}} L_2\right) - \boldsymbol{\gamma}^\top L_{\mathcal{G}} \\ \dot{\boldsymbol{\gamma}} &= L_{\mathcal{G}}\boldsymbol{\lambda} \\ \dot{\boldsymbol{p}} &= -\nabla_{\boldsymbol{p}} L_2 - S_{\boldsymbol{\lambda}}^{\boldsymbol{p}}\nabla_{\boldsymbol{\lambda}} J + S_{\boldsymbol{\lambda}}^{\boldsymbol{p}\,2}\left(\nabla_{\boldsymbol{p}} L_2 + \nabla_{\boldsymbol{\theta}} L_2 - \nabla_{\boldsymbol{\mu}} L_2\right) \\ \dot{\boldsymbol{\theta}} &= -\boldsymbol{\mu}^\top \nabla_{\boldsymbol{\theta}} \boldsymbol{h}(\boldsymbol{p}, \boldsymbol{\theta}) - \tau\nabla_{\boldsymbol{\theta}} \log(-\boldsymbol{g}(\boldsymbol{p}, \boldsymbol{\theta})) \\ \dot{\boldsymbol{\mu}} &= \boldsymbol{h}(\boldsymbol{p}, \boldsymbol{\theta}) \end{aligned} \tag{23}$$

Saddle-point dynamics (23) solves the bi-level optimization problem (8) in a single-level timescale and in continuous-time. The inclusion of network limit constraints can also be included using an additional set of multiplier variables y and a projected dynamic equation $\dot{y} = [\nabla_y L_2]_y^+$ instead of the barrier function used in this work, obtaining then a projected saddle-point dynamics. The next section presents the stability and convergence of this dynamics to the optimal solution asymptotically.

IV. Stability and Convergence Analysis

In this section, it is presented the stability and convergence result. First, we introduce the basic convexity conditions on the Lagrangian to guarantee the convergence of the saddle-point dynamics. As it is noticed in the second-order optimality conditions, in this problem we are minimizing in the variables $\boldsymbol{\lambda}$ and $\boldsymbol{p}$, and maximizing in the Lagrange multipliers $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$. In this sense, we can use the convexity and concavity properties of the Lagrangian functions with respect to each variable. The first-order conditions of convexity for each Lagrangian function considering the bi-level optimization problem are as follows. For the Lagrangian function $L_1(\boldsymbol{\lambda}, \boldsymbol{p}^*(\boldsymbol{\lambda}))$ is

$$\begin{aligned} L_1\big(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \boldsymbol{p}^*(\boldsymbol{\lambda})\big) \geq{}& L_1\big(\boldsymbol{\lambda}^*, \boldsymbol{\gamma}^*, \boldsymbol{p}^*(\boldsymbol{\lambda}^*)\big) \\ &+ \big(\mathcal{D}_\lambda L_1\big(\boldsymbol{\lambda}^*, \boldsymbol{\gamma}^*, \boldsymbol{p}^*(\boldsymbol{\lambda}^*)\big)\big)^\top\big(\boldsymbol{\lambda} - \boldsymbol{\lambda}^*\big), \end{aligned} \tag{24}$$

and for the Lagrangian function $L_2(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu})$ is

$$\begin{aligned} L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}\big) \geq{}& L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}^*, \boldsymbol{\theta}, \boldsymbol{\mu}\big) \\ &+ \big(\nabla_{\boldsymbol{p}} L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}^*, \boldsymbol{\theta}, \boldsymbol{\mu}\big)\big)^\top\big(\boldsymbol{p} - \boldsymbol{p}^*\big), \end{aligned} \tag{25}$$

and

$$\begin{aligned} L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}\big) \geq{}& L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}^*, \boldsymbol{\mu}\big) \\ &+ \big(\nabla_{\boldsymbol{\theta}} L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}^*, \boldsymbol{\mu}\big)\big)^\top\big(\boldsymbol{\theta} - \boldsymbol{\theta}^*\big). \end{aligned} \tag{26}$$

On the other hand, we have the concavity conditions as

$$\begin{aligned} L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}\big) \leq{}& L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\mu}^*\big) \\ &+ \big(\nabla_\mu L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}^*\big)\big)^\top\big(\boldsymbol{\mu} - \boldsymbol{\mu}^*\big). \end{aligned} \tag{27}$$

Considering Assumption 3, the second-order conditions of convexity-concavity follow the second-order optimality conditions in Lemma 2. From the basic notions of saddle-points, a point $(\boldsymbol{\lambda}^*, \boldsymbol{\gamma}^*, \boldsymbol{p}^*(\boldsymbol{\lambda}^*), \boldsymbol{\theta}^*, \boldsymbol{\mu}^*)$ is a local saddle-point of continuously differentiable Lagrangian functions if

$$L_1\big(\boldsymbol{\lambda}^*, \boldsymbol{\gamma}^*, \boldsymbol{p}^*\big(\boldsymbol{\lambda}^*\big)\big) \leq L_1\big(\boldsymbol{\lambda}, \boldsymbol{\gamma}^*, \boldsymbol{p}^*(\boldsymbol{\lambda})\big) \tag{28}$$

and

$$L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}^*, \boldsymbol{\theta}^*, \boldsymbol{\mu}\big) \leq L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}^*, \boldsymbol{\theta}^*, \boldsymbol{\mu}^*\big) \leq L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}^*\big). \tag{29}$$

Now, we have all the elements to establish a stability result. The Local solution concepts introduced in Definition 1 and first and second order necessary and sufficient conditions for bi-level optimization problems presented in Lemma 1 and Lemma 2 are the basis to establish the main stability result for the saddle-point dynamics with predictive-sensitivity conditioning term (23) in the following theorem.

Theorem 1: Consider a local solution $(\boldsymbol{\lambda}^*, \boldsymbol{\gamma}^*, \boldsymbol{p}^*, \boldsymbol{\theta}^*, \boldsymbol{\mu}^*)$ satisfying the sufficient optimality conditions in Lemma 2 and the convexity and concavity conditions. Then, it is a locally asymptotically stable point of the predictive-sensitivity saddle-point dynamics (23)

Proof: Let $(\boldsymbol{\lambda}^*, \boldsymbol{\gamma}^*, \boldsymbol{p}^*, \boldsymbol{\theta}^*, \boldsymbol{\mu}^*)$ be a local solution of the saddle-point dynamics (23), we need to prove that this dynamical system is asymptotically stable. Consider the Lyapunov function $V : X \mapsto \mathbb{R}_{\geq 0}$

$$V(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}) = \frac{1}{2}\Big(||\boldsymbol{\lambda} - \boldsymbol{\lambda}^*||^2 + ||\boldsymbol{\gamma} - \boldsymbol{\gamma}^*||^2 + ||\boldsymbol{p} - \boldsymbol{p}^*||^2 + ||\boldsymbol{\theta} - \boldsymbol{\theta}^*||^2 + ||\boldsymbol{\mu} - \boldsymbol{\mu}^*||^2\Big),$$

The Lie derivative $\mathcal{L}$ of V along the dynamics (23) is as follows

$$\begin{aligned}\mathcal{L}V(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}) = &-\big(\boldsymbol{\lambda} - \boldsymbol{\lambda}^*\big)^\top \mathcal{D}_{\boldsymbol{\lambda}} L_1(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \boldsymbol{p}) \\ &+ \big(\boldsymbol{\gamma} - \boldsymbol{\gamma}^*\big)^\top \nabla_{\boldsymbol{\gamma}} L_1(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \boldsymbol{p}) \\ &- (\boldsymbol{p} - \boldsymbol{p})^\top \nabla_{\boldsymbol{p}} L_2(\boldsymbol{\lambda}, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}) \\ &- \big(\boldsymbol{\theta} - \boldsymbol{\theta}^*\big)^\top \nabla_{\boldsymbol{\theta}} L_2(\boldsymbol{\lambda}, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}) \\ &+ \big(\boldsymbol{\mu} - \boldsymbol{\mu}^*\big)^\top \nabla_{\boldsymbol{\mu}} L_2(\boldsymbol{\lambda}, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}),\end{aligned}$$

following the first-order condition for convexity and concavity (24), (25), (26), and (27) and reordering terms, it is obtained

$$\begin{aligned}\mathcal{L}V(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}) \leq\; & L_1\big(\boldsymbol{\lambda}^*, \boldsymbol{\gamma}, \boldsymbol{p}^*(\boldsymbol{\lambda}^*)\big) - L_1\big(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \boldsymbol{p}^*(\boldsymbol{\lambda})\big) \\ &+ L_1\big(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \boldsymbol{p}^*(\boldsymbol{\lambda})\big) - L_1\big(\boldsymbol{\lambda}, \boldsymbol{\gamma}^*, \boldsymbol{p}^*(\boldsymbol{\lambda})\big) \\ &+ L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}^*, \boldsymbol{\theta}, \boldsymbol{\mu}\big) - L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}\big) \\ &+ L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}^*, \boldsymbol{\mu}\big) - L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}\big) \\ &+ L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}\big) - L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}^*\big),\end{aligned}$$

and from the definition of saddle-point, we verify the semidefiniteness of $\mathcal{L}V(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu})$

$$\begin{aligned}\mathcal{L}V(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}) \leq\; & L_1\big(\boldsymbol{\lambda}^*, \boldsymbol{\gamma}, \boldsymbol{p}^*(\boldsymbol{\lambda}^*)\big) - L_1\big(\boldsymbol{\lambda}, \boldsymbol{\gamma}^*, \boldsymbol{p}^*(\boldsymbol{\lambda}^*)\big) \\ &+ L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}^*, \boldsymbol{\theta}, \boldsymbol{\mu}\big) - L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}\big) \\ &+ L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}^*, \boldsymbol{\mu}\big) - L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}^*\big) \leq 0.\end{aligned}$$

Finally, using the LaSalle invariance principle [32], we need to verify that the only solution belonging to the subspace defined by $\mathcal{L}V = 0$ is the local solution $(\boldsymbol{\lambda}^*, \boldsymbol{\gamma}^*, \boldsymbol{p}^*, \boldsymbol{\theta}^*, \boldsymbol{\mu}^*)$. Notice that $\mathcal{L}V = 0$ implies that

$$\begin{aligned}&L_1\big(\boldsymbol{\lambda}^*, \boldsymbol{\gamma}, \boldsymbol{p}^*(\boldsymbol{\lambda}^*)\big) - L_1\big(\boldsymbol{\lambda}, \boldsymbol{\gamma}^*, \boldsymbol{p}^*(\boldsymbol{\lambda}^*)\big) \\ &+L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}^*, \boldsymbol{\theta}, \boldsymbol{\mu}\big) - L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}\big) \\ &+L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}^*, \boldsymbol{\mu}\big) - L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}^*\big) = 0.\end{aligned}$$

and then

$$\begin{aligned}L_1\big(\boldsymbol{\lambda}^*, \boldsymbol{\gamma}, \boldsymbol{p}^*(\boldsymbol{\lambda}^*)\big) &= L_1\big(\boldsymbol{\lambda}, \boldsymbol{\gamma}^*, \boldsymbol{p}^*(\boldsymbol{\lambda}^*)\big) \\ L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}^*, \boldsymbol{\theta}, \boldsymbol{\mu}\big) &= L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}\big) \\ L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}^*, \boldsymbol{\mu}\big) &= L_2\big(\boldsymbol{\lambda}^*, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\mu}^*\big)\end{aligned}$$

In order to be satisfied for all $t \geq 0$, it is needed that the trajectories $\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}^*$, $\boldsymbol{p}(t) = \boldsymbol{p}^*$, $\boldsymbol{\gamma}(t) = \boldsymbol{\gamma}^*$, $\boldsymbol{\theta}(t) = \boldsymbol{\theta}^*$, and $\boldsymbol{\mu}(t) = \boldsymbol{\mu}^*$ in the invariance set, which is in fact the local solution. Hence, we can conclude that the saddle-point dynamics (23) is asymptotically stable. ■

These stability results assure convergence of the dynamics to the saddle-point and then the solution of the transactive bi-level optimization problem. Another important concern in distributed optimization when the number of DERs is large is the convergence rate. The characterization of the time required by the saddle-point dynamics with the predictive-sensivity conditioning term to converge to the optimal solution depends on the properties of the graph, such as the number of agents and its communication topology. Considering that the proposed saddle-point dynamics is in continuous-time, the solution is also related to the efficiency of the numerical method used to solve the system of differential equations. Hence, when the number of agents is large, acceleration in first-order optimization algorithms has recently emerged as an important research topic [33], and it could be useful to be included into the dynamical system. Some results have analyzed acceleration algorithms as a continuous-time interpretation, indicating that accelerated gradient descent method can be obtained via discretization from a second-order ordinary differential equation [34]. In this way, we could extend the proposed saddle-point dynamics (23) to include a second-order term to accelerate the convergence rate to handle numerous DERs, this is the subject of current research.

In the next section, simulation results are presented to illustrate the capability of the proposed feedback optimization strategy.

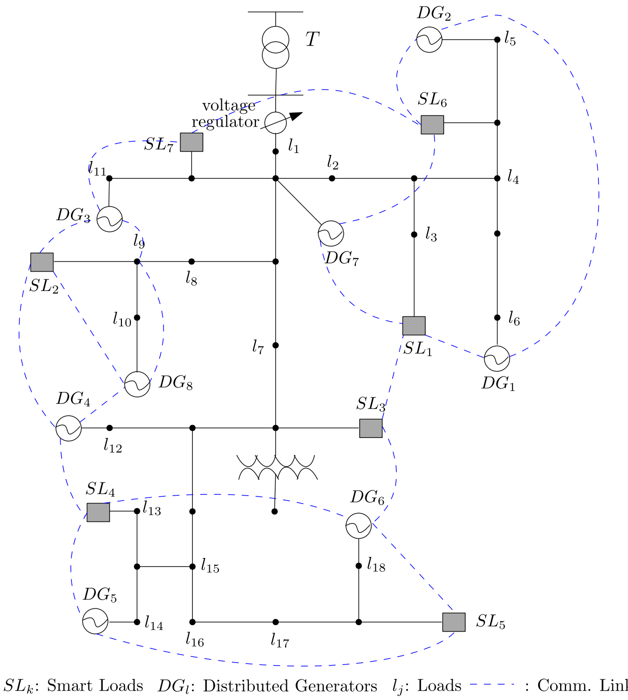

Fig. 3. Distribution network with DERs adapted from an IEEE 37-bus test system.

V. Simulation Results

In this section, a distribution network with DERs is simulated in MATLAB to study the behavior of the proposed approach. We use an adapted IEEE 37-bus test feeder system, as it can be seen in Fig. 3. The systems consist of eight distributed generators with power generation limits and seven consumers able to apply changes in their loads depending on the system state, with a predetermined base load. Furthermore, there is a communication graph network between consumers

TABLE I
SYSTEM PARAMETERS SIMULATION

Generators				
	$\overline{\mathbf{P_{g_i}}}$	$\underline{\mathbf{P_{g_n}}}$	$\beta_{\mathbf{g_n}}$	$\alpha_{\mathbf{g_n}}$
1	100 [MW]	0	2	4
2	150 [MW]	0	2.05	5.88
3	100 [MW]	0	2.1	5.88
4	150 [MW]	0	2.2	5.88
5	100 [MW]	0	2.25	5.88
6	100 [MW]	0	2.3	5.88
7	100 [MW]	0	4	5.88
8	100 [MW]	0	3.2	5.88
Consumers				
	$\overline{\mathbf{P_{d_m}}}$	$\underline{\mathbf{P_{d_m}}}$	$\beta_{\mathbf{d_m}}$	$\rho_{\mathbf{d_m}}$
1	40 [MW]	0	19.2	10.4
2	35 [MW]	0	10	10.4
3	30 [MW]	0	15	10.4
4	25 [MW]	0	11	10.4
5	30 [MW]	0	13	10.4
6	35 [MW]	0	12	22
7	30 [MW]	0	12.15	10.4
Simulation Constants				
	Log barrier term (t)	0.001		
	c	0.0125		
	Simulated iterations	900		
	Stepsize	0.05		
Change of power (iterations)		300	600	

and generators, also shown in Fig. 3 as a dashed line. The data for case studies is presented in Table I. As it is specified in the IEEE 37-bus test feeder system, node 1 is taken as the reference or slack node. Hence, this node will be the reference for the voltage for all nodes, i.e., $V_1 = 1$ p.u.

In contrast to traditional distributed control algorithms, where the focus is in single level problem to obtain the price signal offline based on prediction of load consumption, the proposed framework is able to handle responsive-to-price loads and generators at the same time and in single timescale. Several efficient numerical methods are available to solve the system of differential equations (23). We have used a canonical Runge-Kutta in MATLAB. In order to illustrate the advantages of the proposed framework, we have proposed two simulation case scenarios as follows. Case scenario 1 presents base load variation. Initially, the system has to supply 200 MW from the base load in the distribution network. The agents in the distribution network dynamically interact between them to reach the optimal point, which is a saddle point. Then, at 300 iterations of simulation, an additional load is added to the grid, summing up to 240 MW. Finally, the base load is decreased until 220 MW at 600 iterations. Case scenario 2 was extracted from a real-life instance, where the Uruguayan load profile is considered as follows: the load profile corresponds to the February 1st 2022 between 18.00 to 20.00 h, with set points every 10 minutes. In this range of time, there are several changes in the profile load, where we can see a constant increasing until 19.00 h and decreasing from that time onwards. Therefore, we are controlling the dispatch in one peak of consumption. For this simulation, we assume that the reaction of generators and smart loads can handle the new set points generated by the algorithm.

A. Case Scenario 1

In Fig. 4(a), it is possible to observe the eighth dispatchable generators. They started at a generation set-point that matches the base load, but it is not optimal. It is possible to recognize at the beginning of the simulation how the cheapest generators take most of the power to be generated, meanwhile the more expensive the generator is, less load is assigned to it. This behavior is repeated after the change of the load in iterations 300 and 600. Then, as it is expected, the gap between generators levels is increased proportionally given the need for more power generation. Furthermore, the sum of the generators powers matches perfectly the base load in addition to the load requested by the consumers agents. It is clear that the consumption of the smart loads represents between the 8% and 10% depending on the status of the system.

The consumers behave in a most rich way than generators, and their behavior is shown in Fig. 4(b). They begin searching for the optimal power consumption given the conditions of the system, when they reach the saddle point, they converge to that point. Then, the behavior of the loads can be divided in two. The first stage comprises a base load rise, and then they have to recalculate the optimal consumption. It is predictable that after the load charged the system, the controllable loads slowly decrease until reaching the optimal consumption depending on the pricing and the utility function of each consumer. Furthermore, when the base load in the system is reduced, the consumption of the smart loads increases in a small proportion considering the new price in the system. In addition, in Fig. 4(c) it is possible to see how the price in the nodes behaves with respect to time. It is remarkable that all of them are in a range of 2% of the total price. The differences come from the fact that the price dynamics are affected by the feedback from the power, however the nodal prices are always in a range of 2%, in the worst case scenario when the load is increased. The price dynamics affect as well the smart loads dynamics presenting an inverse effect, i.e., when the price increases, the load demanded by these consumers decreases, and vice-versa. The simultaneous dynamic response of the nodal prices, on one hand, and the generator and smart loads, on the other, shows how the proposed predictive-sensivity conditioning term acting as a feedforward term in the faster system anticipating the dynamics of the slower dynamics improves the convergence rate in comparison with traditional bilevel optimization algorithms in different timescales, which are difficult to solve due to its nested nature and are usually solved by a centralized controller.

Finally, in Fig. 4(d) it is observed how the voltage reference for the node and generator 1 is always maintained in 1 p.u. However, the rest of the nodes present a slight difference with respect to the nominal voltage. These voltages change when it is needed more power flow between lines. In general, it is observed that the difference between nodes is increased, allowing a greater power flow between lines. In this simulation is also illustrated the action of the predictive-sensivity term when the network constraint model is considered in the proposed dynamic algorithm, which leads the distribution system to have small voltages variations while keeping the voltage stability.

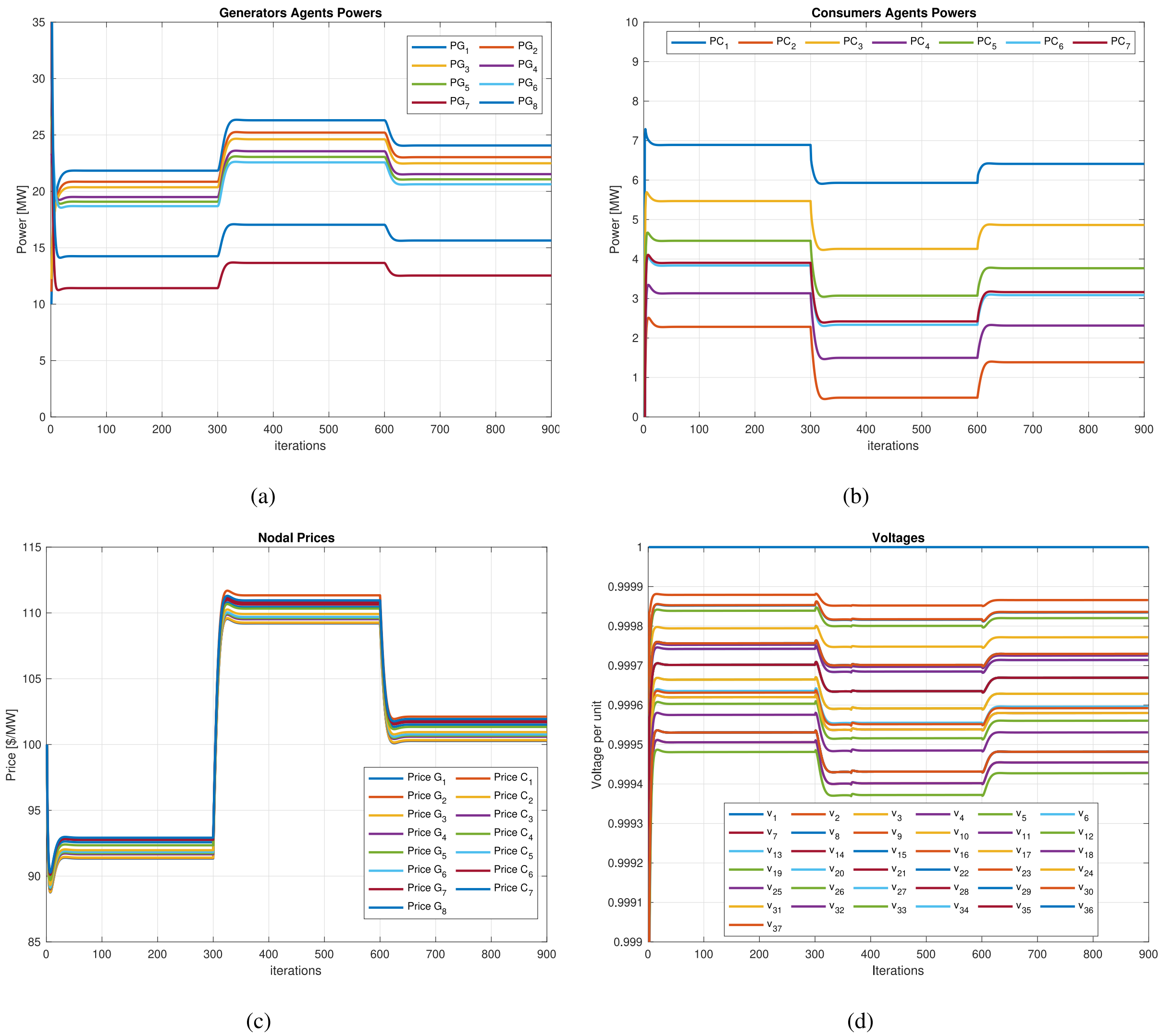

Fig. 4. Simulation results in IEEE 37 node system test case scenario 1 under the distributed transactive control algorithm. (a) Generators dynamic response. (b) Consumers dynamic response. (c) Nodal price behavior. (d) Nodal voltage behavior.

Furthermore, after 600 iterations, it goes back closely to the initial state, when the load is increased again.

B. Case Scenario 2

We evaluate that the proposed algorithm can perform dispatching a real load profile. As in the previous case, we see that the generators agents in modified the power generated proportionally as the grid is demanding power, see Fig. 5(a). It is shown that the most expensive generation unit is the one which has less assigned power, around half the one which is cheaper. Furthermore, we can as well observe the power of the consumer agents vary inversely as the generator power rises or decreases. For instance, it is observed how the smart consumers increase the consumption after the peak has passed, see Fig. 5(b). Basically, this kind of behavior improve the behavior of the load profile and try to flatter the load curve. With respect to the prices, it is shown that when the prices are decreasing, the consumer agents increase the consumption, as it is shown in Fig. 5(c). After every variation of the generation/consumption profile, the price is achieving the optimal stable point, the clearing price. Finally, in Fig. 5(d), a dynamic response similar to the case scenario 1 is observed. The voltage reference for the node and generator 1 is always maintained in 1 p.u. However, the rest of the nodes present a slight difference with respect to the nominal voltage. These voltages change when it is needed more power flow between lines. Nevertheless, the voltage stability is guaranteed since we are considering the network constraints and the dynamic response is fast enough and converge to the optimal equilibrium point.

VI. Conclusion

We have introduced a fully distributed dynamic transactive control to coordinate DERs in a distribution system considering physical network constraints based on saddle-point dynamics with a predictive-sensitivity conditioning term. Theoretical stability result for the asymptotic stability of the saddle-point dynamics with predictive-sensitivity conditioning

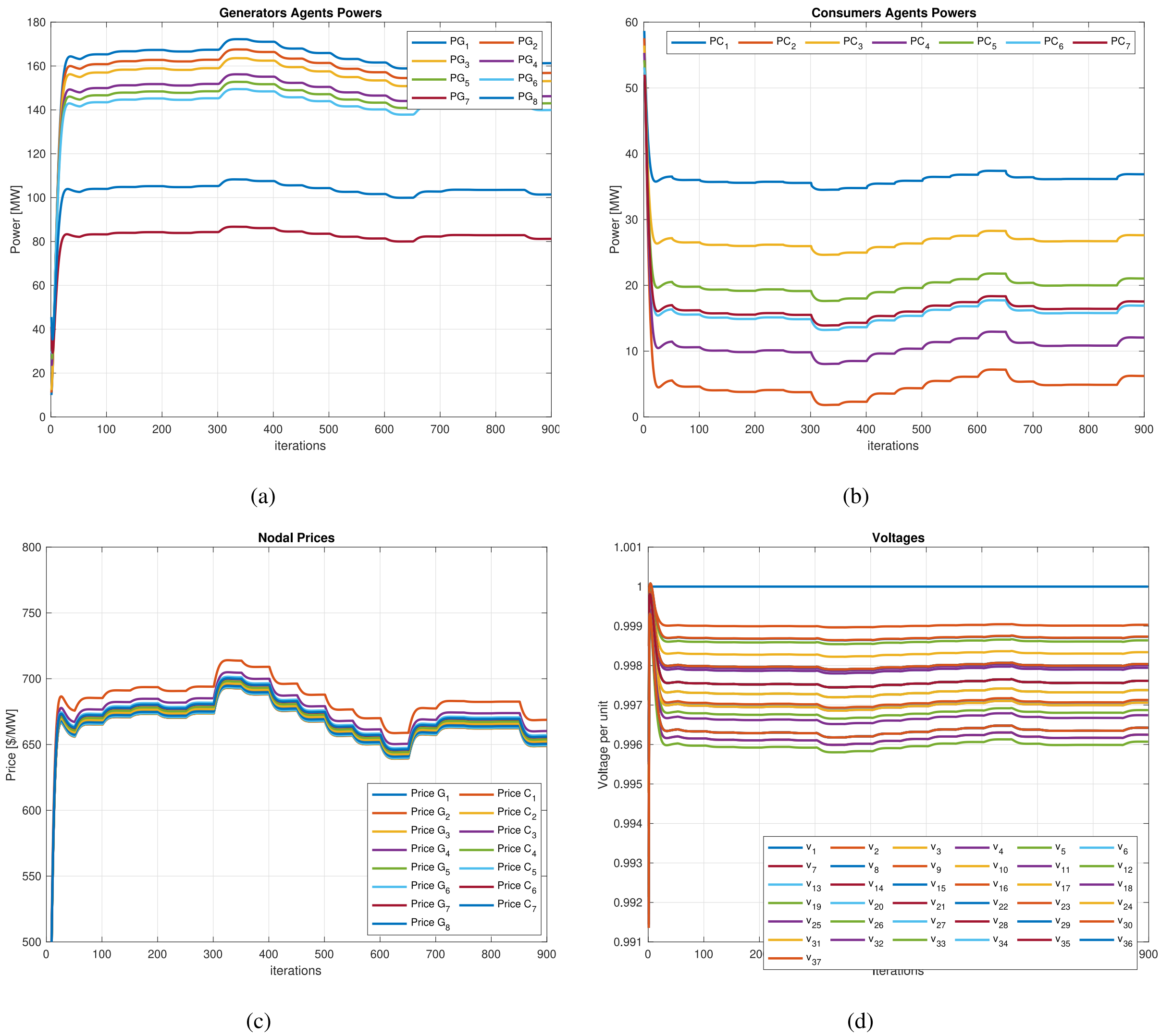

Fig. 5. Simulation results in IEEE 37 node system test case scenario 2 under the proposed distributed transactive control algorithm. (a) Generators dynamic response. (b) Consumers dynamic response. (c) Nodal price behavior. (d) Nodal voltage behavior.

is presented. Simulation results for a distribution network with DERs adopted from an IEEE 37-bus test feeder is implemented and shows that the distributed saddle-point dynamics converges to the optimal solution of the bi-level optimization problem, while keeping the voltage level. As a future work, it would be interesting to include DERs of different nature such as batteries and to analyze the dynamical response of the saddle-point dynamics under communications constraints such as delays or packet losses. On the other hand, it could be important to extend the proposed saddle-point dynamics to include a second-order term to accelerate the convergence rate to handle numerous DERs.

REFERENCES

[1] C. Barreto, E. Mojica-Nava, and N. Quijano, "Incentive mechanisms to prevent efficiency loss of non-profit utilities," *Int. J. Elect. Power Energy Syst.*, vol. 110, pp. 523–535, Sep. 2019.

[2] P. Samadi, H. Mohsenian-Rad, R. Schober, and V. W. S. Wong, "Advanced demand side management for the future smart grid using mechanism design," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1170–1180, Sep. 2012.

[3] K. Kok and S. Widergren, "A society of devices: Integrating intelligent distributed resources with transactive energy," *IEEE Power Energy Mag.*, vol. 14, no. 3, pp. 34–45, May/Jun. 2016.

[4] A. K. Bejestani, A. Annaswamy, and T. Samad, "A hierarchical transactive control architecture for renewables integration in smart grids: Analytical modeling and stability," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 2054–2065, Jul. 2014.

[5] J. Hu, G. Yang, K. Kok, Y. Xue, and H. W. Binder, "Transactive control: A framework for operating power systems characterized by high penetration of distributed energy resources," *J. Modern Power Syst. Clean Energy*, vol. 5, pp. 451–464, May 2017.

[6] R. B. Melton, "Gridwise transactive energy framework," Pacific Northwest Nat. Lab.(PNNL), Richland, WA USA, Rep. PNNL-22946 Ver1.0, 2013.

[7] S. Li, J. Lian, A. J. Conejo, and W. Zhang, "Transactive energy systems: The market-based coordination of distributed energy resources," *IEEE Control Syst. Mag.*, vol. 40, no. 4, pp. 26–52, Aug. 2020.

[8] M. F. Zia, M. Benbouzid, E. Elbouchikhi, S. M. Muyeen, K. Techato, and J. M. Guerrero, "Microgrid transactive energy: Review, architectures, distributed ledger technologies, and market analysis," *IEEE Access*, vol. 8, pp. 19410–19432, 2020.

[9] Y. Cao *et al.*, "Optimal energy management for multi-microgrid under a transactive energy framework with distributionally robust optimization," *IEEE Trans. Smart Grid*, vol. 13, no. 1, pp. 599–612, Jan. 2022.

[10] D. J. Shiltz, M. Cvetković, and A. M. Annaswamy, "An integrated dynamic market mechanism for real-time markets and frequency regulation," *IEEE Trans. Sustain. Energy*, vol. 7, no. 2, pp. 875–885, Apr. 2016.

[11] R. Ambrosio, "Transactive energy systems," *IEEE Electrific. Mag.*, vol. 4, no. 4, pp. 4–7, Dec. 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7725890/

[12] O. Abrishambaf, F. Lezama, P. Faria, and Z. Vale, "Towards transactive energy systems: An analysis on current trends," *Energy Strategy Rev.*, vol. 26, Nov. 2019, Art. no. 100418.

[13] J. Knudsen, J. Hansen, and A. M. Annaswamy, "A dynamic market mechanism for the integration of renewables and demand response," *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 3, pp. 940–955, May 2016.

[14] J. Guerrero, D. Gebbran, S. Mhanna, A. C. Chapman, and G. Verbič, "Towards a transactive energy system for integration of distributed energy resources: Home energy management, distributed optimal power flow, and peer-to-peer energy trading," *Renew. Sustain. Energy Rev.*, vol. 132, Oct. 2020, Art. no. 110000.

[15] M. H. Ullah and J.-D. Park, "Peer-to-peer energy trading in transactive markets considering physical network constraints," *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 3390–3403, Jul. 2021.

[16] E. Mojica-Nava and F. Ruiz, "Stackelberg population dynamics: A predictive-sensitivity approach," *Games*, vol. 12, no. 4, p. 88, 2021.

[17] M. Farivar and S. H. Low, "Branch flow model: Relaxations and convexification—Part I," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2554–2564, Aug. 2013.

[18] S. H. Low, "Convex relaxation of optimal power flow—Part I: Formulations and equivalence," *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 1, pp. 15–27, Mar. 2014.

[19] E. Baron-Prada and E. Mojica-Nava, "A population games transactive control for distributed energy resources," *Int. J. Elect. Power Energy Syst.*, vol. 130, Sep. 2021, Art. no. 106874.

[20] A. Cherukuri, B. Gharesifard, and J. Cortés, "Saddle-point dynamics: Conditions for asymptotic stability of saddle points," *SIAM J. Control Optim.*, vol. 55, no. 1, pp. 486–511, 2017.

[21] K. J. Arrow, H. Azawa, L. Hurwicz, and H. Uzawa, *Studies in Linear and Non-Linear Programming*, vol. 2. Stanford, CA, USA: Stanford Univ. Press, 1958.

[22] T. Kose, "Solutions of saddle value problems by differential equations," *Econometrica J. Econometric Soc.*, vol. 24, no. 1, pp. 59–70, 1956.

[23] J. Wang and N. Elia, "A control perspective for centralized and distributed convex optimization," in *Proc. 50th IEEE Conf. Decis. Control Eur. Control Conf.*, 2011, pp. 3800–3805.

[24] D. Feijer and F. Paganini, "Stability of primal–dual gradient dynamics and applications to network optimization," *Automatica*, vol. 46, no. 12, pp. 1974–1981, 2010.

[25] X. Liu and A. J. Conejo, "Single-level electricity market equilibrium with offers and bids in energy and price," *IEEE Trans. Power Syst.*, vol. 36, no. 5, pp. 4185–4193, Sep. 2021.

[26] A. J. Conejo and C. Ruiz, "Complementarity, not optimization, is the language of markets," *IEEE Open Access J. Power Energy*, vol. 7, pp. 344–353, 2020.

[27] S. Dempe, B. S. Mordukhovich, and A. B. Zemkoho, "Sensitivity analysis for two-level value functions with applications to bilevel programming," *SIAM J. Optim.*, vol. 22, no. 4, pp. 1309–1343, 2012.

[28] J. F. Bard, *Practical Bilevel Optimization: Algorithms and Applications*, vol. 30. New York, NY, USA: Springer, 1998.

[29] S. G. Krantz and H. R. Parks, *The Implicit Function Theorem: History, Theory, and Applications*. New York, NY, USA: Springer, 2012.

[30] D. Bertsekas, A. Nedic, and A. Ozdaglar, *Convex Analysis and Optimization*. Belmont, MA, USA: Athena Sci., 2003.

[31] M. Picallo, S. Bolognani, and F. Dörfler, "Sensitivity-conditioning: Beyond singular perturbation for control design on multiple time scales," *IEEE Trans. Autom. Control*, early access, May 23, 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9779971

[32] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Upper Saddle River, NJ, USA: Patience-Hall, 2002.

[33] Y. Nesterov, "A method of solving a convex programming problem with convergence rate $O(1/k^2)$," *Soviet Math. Doklady*, vol. 27, no. 2, pp. 372–376, 1983.

[34] W. Su, S. Boyd, and E. Candes, "A differential equation for modeling Nesterov's accelerated gradient method: Theory and insights," in *Advances in Neural Information Processing Systems*, vol. 27. Red Hook, NY, USA: Curran, 2014.

Eduardo Mojica-Nava (Member, IEEE) received the B.S. degree in electronics engineering from the Universidad Industrial de Santander, Bucaramanga, Colombia, in 2002, the M.Sc. degree in electronics engineering from the Universidad de Los Andes, Bogotá, Colombia, and the Ph.D. degree in automatique et informatique industrielle from the École des Mines de Nantes, Nantes, France, in co-tutelle with UAndes in 2010. He is currently a Full Professor with the Department of Electrical and Electronics Engineering, Universidad Nacional de Colombia, Bogotá. He has been a Visiting Professor with the University of Mons, Belgium, in 2017 and the Politecnico de Milano, Italy, in 2021. His current research interests include optimization and control of complex networked systems, operator theory method, switched and hybrid systems, and control in smart grids applications.

Fredy Ruiz (Senior Member, IEEE) was born in Facatativa, Colombia. He received the bachelor's and M.Sc. degrees in electronics engineering from Pontificia Universidad Javeriana, Colombia, in 2002 and 2006, respectively, and the Ph.D. degree in computer and control engineering from the Politecnico di Torino, Italy, in 2009. He is currently an Associate Professor with the Politecnico di Milano, Italy. He was an Assistant Professor from 2010 to 2014 and an Associate Professor from 2015 to 2019 with Pontificia Universidad Javeriana, where he also served as the Head of the Electronics Engineering Department from 2014 to 2016. He was a Fulbright visiting Scholar with the University of California at Berkeley in 2013 and a Visiting Professor with the Politecnico di Torino in 2018. His research activity focuses on control and optimization, in particular, the use of data-driven techniques in optimal estimation and control systems design, with applications in smart grids, power electronics, robotics, and biotechnology.

Eder Baron-Prada (Member, IEEE) received the B.S. degree in electrical engineering and the M.S. degree in industrial automation from the Universidad Nacional de Colombia, Bogotá, Colombia, in 2016 and 2019, respectively. In 2022, he joined the Austrian Institute of Technology, Vienna, where he started his joint Ph.D. Project with the Automatic Control Laboratory, Swiss Federal Institute of Technology (ETH) Zurich, Switzerland. His main research interests include optimization and control theory with applications to the power system.