Enabling DER Participation in Frequency Regulation Markets

Priyank Srivastava  Chin-Yao Chang  Jorge Cortés

Abstract—Distributed energy resources (DERs) are playing an increasing role in ancillary services for the bulk grid, particularly in frequency regulation. In this paper, we propose a framework for collections of DERs, combined to form microgrids and controlled by aggregators, to participate in frequency regulation markets. Our approach covers both the identification of bids for the market clearing stage and the mechanisms for the real-time allocation of the regulation signal. The proposed framework is hierarchical, consisting of a top layer and a bottom layer. The top layer consists of the aggregators communicating in a distributed fashion to optimally disaggregate the regulation signal requested by the system operator. The bottom layer consists of the DERs inside each microgrid whose power levels are adjusted so that the tie line power matches the output of the corresponding aggregator in the top layer. The coordination at the top layer requires the knowledge of cost functions, ramp rates and capacity bounds of the aggregators. We develop meaningful abstractions for these quantities respecting the power flow constraints and taking into account the load uncertainties, and propose a provably correct distributed algorithm for optimal disaggregation of regulation signal amongst the microgrids.

1 INTRODUCTION

Electric power systems require the generation and load to be equal at all times. Any discrepancy between the two leads to the deviation of the frequency from its nominal value. This deviation of the frequency leads to many undesirable scenarios. Based on measurements of the frequency deviation, the system operator computes the automatic generation control (AGC) signal as the feedback frequency control to the power system, which appears as the total active power adjustment. Traditionally, frequency regulation services have been provided by individual energy resources, such as coal generation plants or gas turbines. Recently, there has been a trend towards the integration of more DERs into the system to provide these services while reducing thermal and CO₂ emissions. Such integration leads to higher uncertainty in the bulk grid. At the same time, as most DERs are inertialess, they can be effective for frequency regulation due to their high ramp rates. DERs are limited in size and might not meet the minimum size criteria specified by system operators to participate in the frequency regulation market. To address these challenges, the vision is to integrate groups of DERs through distributed energy resource providers (DERPs), or aggregators, which would act as virtual power plants (VPPs) and would be communicating with the system operator. These aggregators do not necessarily own the DERs, they just coordinate their responses. This architecture, illustrated in Figure 1, has been proposed by the California ISO (CAISO) to offer aggregators of DERs the opportunity to sell into its marketplace. The recent Order No. 2222 by the U.S. Federal Energy Regulatory Commission (FERC) also enables aggregators to participate in the energy markets and requires all Regional Transmission Organizations (RTOs) to revise their tariffs to establish DERs as a category of market participant. Using aggregators not only solves the problem of limited capabilities of DERs but also enables the system operator to interact with much fewer entities. This paper is motivated by the need to address the challenges to carry out the vision described above.

Literature Review: Order No. 755 issued by the FERC requires RTOs to compensate energy resources based on the actual frequency regulation provided. The payment to resources comprises of two parts, the capacity and performance payments. The capacity payment compensates resources for their provision of regulation capacity. The performance payment reflects the accuracy of the tracking of the allocated regulation signal. The work describes how different RTOs across the United States have implemented
FERC Order 755 for participation of resources in frequency regulation market. In the literature on power networks and smart grid, some works have considered the possibility of obtaining frequency regulation services from collections of homogeneous loads such as electric vehicles (EVs) and thermostatically controlled loads (TCLs), cf. [6]–[8]. The work [9] presents a method to model flexible loads as a virtual battery for providing frequency regulation. [10] proposes the use of aggregators to integrate heterogeneous loads such as heat pumps, supermarket refrigerators and batteries present in industrial buildings to provide frequency regulation. The work [11] describes the challenges that need to be overcome for providing frequency regulation by DERs for some European countries. The work [12] provides a framework to emulate virtual power plants (VPPs) via aggregations of DERs and provide regulation services taking into account the power flow constraints. [13] provides a dispatch strategy for an aggregator of ON/OFF devices to provide frequency regulation. In [14]–[16], work has been done in the context of microgrids to design mechanisms for optimally allocating a given signal among the DERs within the microgrid. [17] proposes a distributed algorithm to minimize the aggregated cost while satisfying the local constraints and collective demand constraint at the aggregator. However, the aforementioned works assume that the allocated signal from the RTO is available to the aggregator. [18] applies machine learning to forecast the power capacity of VPPs. The work [19] provides a framework for optimal bidding and dispatch of multiple VPPs. Here, we focus on (i) participation of microgrids in frequency regulation markets operated by the RTO through the identification of appropriate bids and (ii) the coordination among RTO and aggregators to efficiently dis-aggregate the regulation signal amongst the aggregators. The actual tracking performance within the microgrid would depend on the physical condition of the resources. In [20], we provide some results for this via experiments carried out on the University of California, San Diego (UCSD) microgrid.

Statement of Contributions: We propose a hierarchical framework for the participation of microgrids in the frequency regulation market. We start by briefly reviewing the current practice of frequency regulation from individual resources, consisting of three stages: (i) market clearance, (ii) disaggregation of the regulation signal and (iii) real-time tracking of the regulation signal. Our first contribution is the identification of the limitations of current practice and the challenges that need to be overcome for integration of microgrids. Our second contribution is the identification of abstractions for the capacity, cost of generation, and ramp rates of a microgrid as a combination of the individual energy resources that compose it, along with a formal description of its convexity and monotonicity properties. Building on our preliminary work [1], here we extend our abstractions to the case when the loads inside the microgrid do not remain constant for the regulation period. Equipped with these abstractions, a microgrid can submit bids to participate in the market clearance stage. Our third contribution is the reformulation of the RTO-DERP coordination problem to optimally disaggregate regulation signal amongst the microgrids and accompanying design of an algorithmic solution. Our proposed reformulation ensures feasibility. The proposed algorithm is distributed over directed graphs with only one aggregator needing to know the required regulation, and is guaranteed to asymptotic converge to the desired optimizers. We conclude with simulation results based on the proposed abstractions of capacities, cost, and ramp rate and the RTO-DERP coordination algorithm on a reduced-order model of the University of California, San Diego (UCSD) microgrid.

2 Preliminaries

In this section, we present notational conventions and review some basic concepts.

Notation: Let \( \mathbb{C}, \mathbb{R}, \mathbb{R}_{\geq 0}, \) and \( \mathbb{Z} \) be the set of complex, real, non-negative real and integer numbers, respectively. For a set \( |X| \), we let \( |X| \) denote its cardinality and \( 1 \) and \( 0 \) denote the vectors of all ones and all zeros of appropriate dimension, respectively. We use \(|x|\) to denote the absolute value of \( x \), \(|x|+\) to denote \( \max\{x,0\} \) and \(|x|_a\) to denote \( |x|+\) if \( a > 0 \) and \( 0 \) if \( a \leq 0 \). If \( x \) is a vector, these functions are applied elementwise. For a matrix \( A \), its \( i \)th row and transpose are denoted by \( A_i \) and \( A^\top \), respectively. We denote the gradient of a differentiable real-valued function \( f: \mathbb{R}^n \to \mathbb{R} \) by \( \nabla f \).

Graph Theory: We let \( G = (\mathcal{V}, \mathcal{E}, A) \) denote a directed graph, with \( \mathcal{V} \) as the set of vertices (or nodes) and \( \mathcal{E} \subseteq \mathcal{V} \times \mathcal{V} \) as the set of edges. \((i, j) \in \mathcal{E} \) iff there is an edge from node \( i \) to \( j \). We let \(|\mathcal{V}| = n\) and \(|\mathcal{E}| = m\). A path is an ordered sequence of vertices such that any pair of vertices that appear consecutively is an edge. A loop is a path in which the first and last vertices are same and none of the other vertices is repeated. A graph is strongly connected if there is a path between any two distinct vertices. A tree is a graph whose underlying undirected graph does not have any loops and is connected. The adjacency matrix \( A \in \mathbb{R}^{n \times n} \) of \( G \) is defined such that \( A_{ij} > 0 \) if the edge \((i, j) \in \mathcal{E} \) and \( 0 \), otherwise. The out-degree and in-degree of a node \( i \) are respectively, the number of outgoing edges from and incoming edges to \( i \). The weighted out-degree and the weighted in-degree of a node \( i \) are given by \( d_{out}^w = \sum_{j=1}^n A_{ij} \) and \( d_{in}^w = \sum_{j=1}^n A_{ji} \), respectively. The weighted out-degree matrix \( D_{out}^w \in \mathbb{R}^{n \times n} \) and the weighted in-degree matrix \( D_{in}^w \in \mathbb{R}^{n \times n} \) are the diagonal matrices with \( D_{out}^w = d_{out}^w \) and \( D_{in}^w = d_{in}^w \). A graph is weight-balanced if \( D_{out}^w = D_{in}^w \). The Laplacian matrix \( L \in \mathbb{R}^{n \times n} \) is defined as \( L = D_{out}^w - A \). 0 is a simple eigenvalue of \( L \) with eigenvector \( 1 \) iff \( G \) is strongly connected, and \( 1^\top L = 0 \) iff \( G \) is weight-balanced. The incidence matrix \( M \in \mathbb{R}^{n \times m} \) is defined such that \( M_{ij} = 1 \) if the edge \( j \) leaves vertex \( i \), \(-1 \) if it enters vertex \( i \), and \( 0 \) otherwise. Note that every column of \( M \) has only two non-zero entries and \( 1^\top M = 0 \). The fundamental loop matrix \( N \in \mathbb{R}^{n \times (m-n+1)} \) of a graph has \( N_{ij} \) as 1 \(-1\) respectively if the \( i \)th edge has the same (opposite, respectively) orientation as the \( j \)th loop, and \( N_{ij} = 0 \) if edge \( i \) is not part of loop \( j \). We use \( P_{ref} \in \mathbb{R}^{(n-1) \times n} \) to denote the path matrix of a tree with reference vertex \( ref \): the \( ij \)th entry of the path matrix is \(+1/-1\) if edge \( j \) is in the directed path from \( i \) to \( ref \) and has the same/opposite orientation as this path, and \( 0 \) otherwise.

Probability Theory: Given an event \( E \), we let \( E^c \) denote its complement and \( \Pr(E) \) its probability. Given a normally distributed random variable \( \zeta \sim \mathcal{N}(\mu, \sigma) \) with mean \( \mu \) and
variance $\sigma$, the probability $\Pr(\zeta \leq x)$ of $\zeta$ being less than or equal to $x$ is denoted
\[
\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(z-x)^2}{2}} dz.
\]

The error function $\text{erf}$, defined as
\[
\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-u^2} du,
\]
denotes the probability of a normal random variable with mean 0 and variance 1/2 being in the interval $[-x, x]$. For a normal random variable with mean 0 and variance 1/2, the functions $\Phi$ and $\text{erf}$ are related by
\[
\Phi(x) = \frac{1}{2} \left(1 + \text{erf} \left( \frac{x}{\sqrt{2}} \right) \right).
\]

**Dynamic Average Consensus:** Consider a network of $n \in \mathbb{Z}_{>1}$ agents communicating over a strongly connected weight balanced directed graph $G$. Each agent has a state $z_i \in \mathbb{R}$ and an input signal $u_i : \mathbb{R} \to \mathbb{R}$. Dynamic average consensus aims at making each agent track the average input $\frac{1}{n} \sum_{i=1}^{n} u_i(t)$ asymptotically. Formally, we employ the dynamics given by
\[
\dot{z} = \dot{u} - \nu(z - u) - \beta L z - v,
\]
\[
\dot{v} = \nu \beta L z,
\]
where $L \in \mathbb{R}^{n \times n}$ is the Laplacian of $G$ and $\nu, \beta > 0$ are the design parameters. If the algorithm is initialized with $1^T v(0) = 0$, then the steady-state error between the state $z_i$ of each agent $i \in \{1, \ldots, n\}$ and the average signal $\frac{1}{n} \sum_{i=1}^{n} u_i$ is bounded, and goes to zero if $\dot{u} \to 0$, cf. [21, Theorem 4.1].

## 3 Frequency Regulation with Microgrids

We are interested in coordinating power aggregators to collectively provide frequency regulation. An aggregator is a virtual entity that aggregates the actions of a group of distributed energy resources to act as a single whole. In this paper, we identify an aggregator with a microgrid, but in general it may correspond to other entities (such as, for instance, a collection of microgrids).

### 3.1 Review of Current Practice

The frequency regulation market is operated by an RTO to regulate the system frequency at its nominal value. To achieve this, the RTO coordinates the response of participating energy resources in a centralized fashion to assign the regulation signal and restore the power balance of the grid. Different RTOs follow slightly different procedures for the frequency regulation markets. The procedure followed by CAISO has the following stages, see e.g., [3].

**[CP1]: Market clearance.** All participating resources submit their capacity bids, capacity price bids, and mileage price bids to the RTO. Capacity bids are the maximum amount of regulation (up or down) that the resource can provide. Capacity price bids are the unit price of providing these regulations. Mileage is the sum of the absolute change in AGC set points, which corresponds to the summation of the vertical lines in Figure 2. The mileage price bid is the cost for unit change in regulation. Typically, expected mileages are calculated from historical data and resources do not submit mileage bids. The RTO clears the market with a capacity and mileage price that is uniform across the resources, and sends each resource its capacity and mileage allocation. This off-line process happens only once per regulation event.

**[CP2]: Allocation of regulation signal to each resource.** The RTO sends the regulation set points to each of the procured energy resources every 2-4 seconds for the entire regulation period, which is usually 10-15 minutes. The regulation set points are computed from the AGC signal in real time in proportion to the procured mileage of each resource. In case the assigned capacity of a resource is violated, the overshoot power is redistributed to the other resources in proportion to their assigned mileages.

**[CP3]: Real-time tracking of regulation signal.** Once the regulation set points have been assigned, the resources need to track them in real time.

Payment to the resources comprises of two components, capacity payment and mileage payment. The capacity payment is done based on the assigned capacity in [CP1] while the mileage payment is done based on the actual mileage provided which reflects the performance of the resources while tracking the assigned signal in [CP3].

### 3.1.1 Limitations of Current Practice

The centralized way of assigning the set points to the resources in [CP2] relies on the fixed number of resources with fixed generation capacities procured in [CP1], which are available for the entire regulation period. This is problematic in the context of aggregators, as they are subject to variabilities and uncertainties associated with the DERs inside them. Even if the DERs inside the microgrid participating stay available for the entire regulation period, the users inside the microgrid can change their power consumption, which in turn leads to changes in the effective regulation capacity. Furthermore, in current practice, there is no direct consideration of the operational costs of the resources, which may result in suboptimal power allocation. Instead, we argue that the assignment of the regulation signal should be done, at each time step, in a way that optimizes the aggregate cost functions of the resources and takes into account their (possibly...
dynamic) operational limits. We refer to this approach as the RTO-DERP coordination problem. This idea has also been pointed out in the past by CAISO for traditional energy resources, cf. [23]. The lack of robustness and the information sharing requirements of centralized schemes motivate the investigation of distributed schemes to solve the RTO-DERP coordination problem.

3.1.2 Challenges for Frequency Regulation from Microgrids
Here we describe the challenges specific to microgrid participation in frequency regulation markets. First, note that solving the RTO-DERP coordination problem with microgrids requires the identification, or rather the abstraction, of aggregate cost functions and regulation capacity bounds based on the cost functions and flexibilities of their DERs. Second, the determination of capacity bids requires taking into account the uncertainties associated with the microgrids. There is a need to calculate bids for each regulation interval, as they might need to considerably change from one interval to the next. Even within a regulation interval itself, the power level of the uncontrollable nodes might vary significantly. Third, the determination of mileage bids has to take into account the dependency of ramp rates on the composition and participation of the individual DERs. The current method of calculating expected mileages in [CP1] makes sense for conventional resources as their ramp rates are fixed and historical data provides reliable accuracy. However, in the case of microgrids, individual resources may be changing over time and as a result, ramp rates might not remain constant. Also, the performance of participating resources for one regulation period might be substantially different from those over a different regulation period.

3.2 Problem Statement
Consider $N$ microgrids, each controlled by an aggregator. To enable microgrid participation in the frequency regulation market, we focus on [CP1] and [CP2]. Based on the discussion in Section 3.1 we formalize the following problems.

[P1]: Meaningful abstractions for the microgrid. To enable the submission of bids in [CP1], each aggregator needs to quantify the maximum up/down regulation capacity that the microgrid can provide, the unit cost of providing such regulation, and the ramp rate at which the microgrid can change its power level. Our first goal is therefore to provide meaningful abstractions for these objects, and cost functions and ramp rate functions of the microgrids for [P2] below, a problem we tackle in Section 5.

[P2]: RTO-DERP distributed coordination. The RTO-DERP coordination problem for computing the set points for each resource advocated for [CP2] consists of an economic dispatch problem with ramp rate constraints at every instant of the regulation interval. Formally, for $x$, regulation at a given time instant, we have

$$\min_x f(x) = \sum_{i=1}^{N} f_i(x_i)$$

subject to

$$\sum_{i=1}^{N} x_i = x_r$$

$$\pi_l \leq x_i \leq \pi_u \quad \forall i$$

where $x \in \mathbb{R}^N$ is the vector of regulation power from the microgrids, $f_i(x_i)$ is the cost of $x_i$ regulation for microgrid $i$, $\pi_l$ and $\pi_u$ are the lower and upper bounds of regulation for microgrid $i$ which are bounded by the solutions of [P1] and determined by [CP1] for a specific regulation period, $x_r$ is the regulation that the microgrid $i$ was providing at the previous instant, and $R_i(x_i)$ is the ramp rate of the microgrid when it is providing regulation $x_i$. Because of the ramp constraints present in [P1], this problem might not be always feasible (since mileage requirements set by the RTO while clearing the market in [CP1] capture the average mileage required, and not the extreme cases). In such cases, we want to minimize the error between the procured regulation and the required one. We tackle these in Section 5.

4 Microgrid Abstractions
Consider a microgrid with $n \in \mathbb{Z}_{>1}$ buses, described by $G_m = (V, E, \Lambda)$. Without loss of generality, we assume that the first bus is connected to the bulk grid through a tie line. We partition the remaining set of buses in $V_g \cup V_l$, where $V_g$ is the set of controllable nodes and $V_l$ is the set of uncontrollable nodes (or loads). Let $n = |V|$, $n_g = |V_g|$, $n_l = |V_l|$ and $m = |E|$. We assume that the network, inverter filter, and voltage controller dynamics are fast enough so that we can model them as power injections with no dynamics [23], [24]. We adopt the convention that the value of the power injection is negative if the device consumes power and vice versa. The power level of each controllable node $p \in V_g$ is denoted by $g_p$, with $g_p^0$ denoting the baseline generation/consumption. The power level of each uncontrollable node $q \in V_l$ is denoted by $l_q$. We denote the incoming power through the tie line by $P$ and its baseline value by $P^0$. When the microgrid provides frequency regulation, the value of the tie line power $P$ is

$$P = P^0 + x,$$

where $x$ is the allocated AGC signal. Note that since we model $P$ as the incoming power from bulk grid, $x$ would be negative when the microgrid is providing up regulation.

Next, let us discuss the power flow equations inside the microgrid following [25]. Let $V_i \in \mathbb{C}$ be the voltage at the $i$th bus, with $|V_i|$ as the magnitude and $\theta_i$ as the phase angle. The power injection $P_i$ at node $i \in V$ is given by

$$P_i = \text{real}(V_i^*I_i^*) = \text{real}\left(V_i\left(\sum_{k=1}^{n} Y_{ik}V_k\right)^*\right)$$

$$= \sum_{k=1}^{n} |V_i||V_k|(G_{ik}\cos(\theta_{ik}) + B_{ik}\sin(\theta_{ik})),$$

where $\theta_{ik} = \theta_i - \theta_k$ and $Y_{ik}, G_{ik}$ and $B_{ik}$ are admittances, conductances and susceptances between bus $i$ and $k$, respectively. Let $\omega_{ik} = |V_i||V_k|(G_{ik}\cos(\theta_{ik}) + B_{ik}\sin(\theta_{ik}))$ represent the line flow from bus $i$ to bus $k$. In this expression, the term $G_{ik}\cos(\theta_{ik})$ corresponds to the power loss part. We make the following simplifying assumptions.

Assumption 1. (Approximate constant voltage): The voltage magnitude of every bus is approximately 1 p.u.
Assumption 2. (Lossless network): The losses in the power flow are negligible, i.e., \( G_{ik} \approx 0 \) for every \( \{i, k\} \in \mathcal{E} \).

Under Assumptions 1 and 2, the power flow from bus \( i \) to bus \( k \) is \( \omega_{ik} = B_{ik} \sin \theta_{ik} \) and the power injection \( \mathcal{P}_i \) at node \( i \) reduces to \( \mathcal{P}_i = \sum_{k=1}^n \omega_{ik} = \sum_{k=1}^n B_{ik} \sin \theta_{ik} \). For the microgrid, we therefore have

\[
\begin{bmatrix}
P \\
g^T \\
-I^T
\end{bmatrix}^T = M B \sin(M^T \theta),
\]

(3)

where \( g \in \mathbb{R}^n \) and \( l \in \mathbb{R}^m \) are the vectors of controllable and uncontrollable nodes, respectively, \( M \in \mathbb{R}^{n \times m} \) is the incidence matrix of the graph, \( B \in \mathbb{R}^{m \times m} \) is the diagonal matrix of absolute line susceptances and \( \theta \in \mathbb{R}^n \) is the vector of phase angles. Additionally, there is a constraint that the maximum phase angle difference between any two interconnected buses should be bounded by \( \Theta \in \mathbb{R}_{\geq 0} \), i.e.,

\[
|\theta_i - \theta_k| \leq \Theta.
\]

(4)

To avoid dealing with the nonlinearity in (3), we make the following assumption for simplicity.

Assumption 3. (No common edge between loops): \( \mathcal{G}_m \) is a graph with non-overlapping loops.

This assumption allows to rewrite the power flow equations in terms of the line flows as follows,

\[
\begin{bmatrix}
P \\
g^T \\
-I^T
\end{bmatrix}^T = M \omega,
\]

(5a)

\[
|\omega| \leq \omega.
\]

(5b)

Here \( \omega \in \mathbb{R}^m \) is the vector of line flows and \( \omega \in \mathbb{R}^m \) is the vector of maximum permissible flows. Using the fact that the columns of the fundamental loop matrix forms a basis for the null space of the incidence matrix, cf. [26 Theorem 4-6], we rewrite (5) as

\[
\begin{bmatrix}
1^T l - 1^T g \\
g \\
l
\end{bmatrix}^T + N \gamma |\leq \omega,
\]

(6)

where \( M^+ \) denotes the Moore-Penrose pseudoinverse of \( M \), \( N \in \mathbb{R}^{m \times (m-n+1)} \) is the fundamental loop matrix of \( \mathcal{G}_m \), and \( \gamma \in \mathbb{R}^{m-n+1} \).

4.1 Capacity Bounds

In order for a microgrid to find the maximum up (or down) regulation that it can provide, it needs to solve an optimization problem. For up regulation, the power consumption of the microgrid is less than the baseline power. Since the baseline power is constant for the regulation period, computing the capacity is equivalent to minimizing \( P \) while satisfying the power flow constraints. If the power level of uncontrollable nodes is constant for the entire regulation period, then the problem reads as

\[
\begin{align*}
\min_{g, \omega} & \quad P \\
\text{s.t.} & \quad \begin{bmatrix} P & g^T & -I^T \end{bmatrix}^T = M \omega \\
& \quad g \leq g \leq \bar{g}, \quad |\omega| \leq \omega,
\end{align*}
\]

(7)

where \( g \) and \( \bar{g} \) are the vectors of minimum and maximum possible power levels of controllable nodes, respectively. If \( \bar{P} \) denotes the solution of (7), then the maximum up regulation is \( \bar{P} = P - P^0 \). The maximum down regulation \( \bar{P} \) can be obtained solving a similar maximization problem.

The formulation (4) assumes that the power level of the uncontrollable nodes remains constant, and therefore does not take into account the varying nature of the loads. In practice, this makes sense for a specific regulation instant, and would rarely be the case for the whole regulation period. Instead, a more robust way of calculating the capacity bounds that the aggregator can use in bidding for the whole regulation period is to account for worst-case scenarios, i.e., taking the expected maximum value for the uncontrollable nodes while computing the maximum up regulation. Although robust to variations in the uncontrollable nodes’ powers, this way of computing capacity bounds might be too conservative and, in fact, might prohibit the microgrid from participating in the regulation market at all.

As an alternative, we propose a reformulation of problem (7) based on chance constraints. Using (6), we rewrite the optimization problem (7) as

\[
\begin{align*}
\min_{g, \gamma, t} & \quad t \\
\text{s.t.} & \quad t \geq 1^T l - 1^T g \\
& \quad |M^+ \begin{bmatrix} 1^T l - 1^T g \\
g \\
l
\end{bmatrix} + N \gamma | \leq \omega \\
& \quad g \leq g \leq \bar{g}.
\end{align*}
\]

(8)

Assume that a probability distribution describing the power levels of uncontrollable nodes at any instant of the regulation period is available. To account for load variability, we instead consider the following chance-constrained optimization

\[
\begin{align*}
\min_{g, \gamma, t} & \quad t \\
\text{s.t.} & \quad \Pr(t \geq 1^T l - 1^T g) \geq 1 - \epsilon' \\
& \quad \Pr(|M^+ \begin{bmatrix} 1^T l - 1^T g \\
g \\
l
\end{bmatrix} + N \gamma | \leq \omega) \geq 1 - \epsilon \quad \forall j \\
& \quad g \leq g \leq \bar{g}.
\end{align*}
\]

(9)

where \( \epsilon', \epsilon \in [0, 1] \). In this formulation, each flow constraint can be violated, with a probability no more than \( \epsilon' \). The following result, whose proof is given in the Appendix, shows that the chance-constrained optimization (9) can be solved via a deterministic linear program if the loads are normally distributed.

Lemma 4.1. (Capacity bounds for variable loads via deterministic optimization): Assume the loads are distributed normally with mean \( \bar{l} \) and variance \( V_l \). Then, the solution of the deterministic linear program

\[
\begin{align*}
\min_{g, \omega} & \quad P \\
\text{s.t.} & \quad \begin{bmatrix} 1^T \bar{l} - 1^T g - t \leq \sqrt{2} \operatorname{erf}^{-1}(2 \epsilon - 1)(1^T V_l 1)^{1/2} \\
& \quad |(M_1 1^T - M_3) l + (M_2 - M_1 1^T) g + N \gamma | \leq \omega \\
& \quad g \leq g \leq \bar{g}
\end{align*}
\]

(10)

where \( M^+ = [M_1 \quad M_2 \quad M_3] \) with \( M_1 \in \mathbb{R}^m, M_2 \in \mathbb{R}^{m \times n_g} \) and \( M_3 \in \mathbb{R}^{m \times n_{\omega}}, \omega' = \omega + K \) and
\[
K_j = \sqrt{2} \text{erf}^{-1}(\epsilon - 1),
\]
\[
\cdot (\langle M_{1j} 1^T - M_{3j} \rangle \mathbf{v}(\langle M_{1j} 1^T - M_{3j} \rangle)^T)^{1/2},
\]
is a solution of problem (9).

![Fig. 3. Regulation capacities for different instantiations of the reduced-order UCSD microgrid. The 12 microgrids are divided into 4 groups, with constant mean value of loads and baseline generation across each group. Within each group, for the first scenario, the variance is 0 and for the remaining two scenarios, it is given by the diagonal matrix which is 0.25 times the squared value of mean loads. Values of \(\epsilon, \epsilon'\) for the second and third scenarios are \(10^{-1}, 4.2 \times 10^{-5}\) and \(2 \times 10^{-1}, 8.4 \times 10^{-5}\), respectively.](image)

We use Lemma 4.1 to compute in Figure 3 the maximum up and down regulation for several microgrids modeled after the reduced-order UCSD microgrid described later in Section 6. The microgrids are divided into 4 groups, each with a different average value of baseline generation and mean load for the UCSD model. Within each group, we consider 3 different scenarios, one with constant load and the other two with varying loads (generated using the same normal distribution with variance equal to the diagonal matrix given by 0.25 times the squared value of mean loads) and different confident values (\(\epsilon, \epsilon' = 10^{-1}, 4.2 \times 10^{-5}\) and \(2 \times 10^{-1}, 8.4 \times 10^{-5}\), respectively). One can see in Figure 3 that the capacity bounds increase with \(\epsilon', \epsilon\), which is in agreement with the fact that larger values of these correspond to lower probability of satisfying the constraints.

Note that the probabilistic capacity bounds identified above and obtained after solving (10) are good only for the bidding in [CP1]. The actual regulation bounds at a given regulation instant still depend on the load at that instant.

### 4.2 Ramp Rate Function

In the following we discuss how to compute the ramp up rate for the microgrid (the discussion for ramp down rate is analogous). If there were no constraints on the power flows, then the ramp rate of the microgrid would be the summation of ramp rates of all the controllable nodes. However, the presence of flow constraints may prevent every controllable node from ramping at its full capacity and as such, the ramp rate is a function that depends on the operating point of the controllable nodes. Let \(F_g = \{g \in \mathbb{R}^{n_g} \mid \exists \omega \in \mathbb{R}^m \text{ satisfying (5)}\}\) denote the set of feasible operating points for controllable nodes. If the power levels of the uncontrollable nodes are constant, then the ramp up rate, \(R: F_g \to \mathbb{R}_{\geq 0}\), is formally given by

\[
\max_{\Delta g, \Delta \omega} \mathbf{1}^T \Delta g
\]

s.t.

\[
\left\lfloor \left(P^T - \mathbf{1}^T \Delta g\right) \left(g + \Delta g\right) - l^T \right\rfloor = M(g + \Delta \omega)
\]

\[
\Delta g \leq r,
\]

\[
|\omega + \Delta \omega| \leq \Delta,
\]

where \(r \in \mathbb{R}^{n_g}\) is the vector whose component \(r_p\) is the nominal ramping capacity of the controllable node \(p\), and \(\omega + \Delta \omega\) is the vector of line flows corresponding to the operating point \(g + \Delta g\).

If the power levels of the uncontrollable nodes are variable, we use chance-constraints as in the case of capacity bounds and the ramp up rate \(R\) is given by

\[
\max_{\Delta g, \gamma} \mathbf{1}^T \Delta g
\]

s.t.

\[
\text{Pr}\left|\mathbb{M}^T \left(1^T l - 1^T (g + \Delta g)\right) + N \gamma\right|_j \leq \Delta_j \geq 1 - \epsilon \forall j
\]

The following result, whose proof is similar to proof of Lemma 4.1 and omitted to avoid repetition, converts the chance-constrained optimization (12) into a deterministic linear program if the loads are normally distributed.

**Lemma 4.2.** (Ramp rate for variable loads via deterministic optimization): Assume the loads are distributed normally with mean \(l\) and variance \(V_l\). Then, the solution of the deterministic linear program

\[
\max_{\Delta g, \gamma} \mathbf{1}^T \Delta g
\]

s.t.

\[
\left|\mathbb{M}^T \left(1^T l - 1^T (g + \Delta g)\right) + N \gamma\right|_j \leq \Delta_j
\]

\[
\Delta g \leq r,
\]

where \(\mathbb{M}, l, V_l\) and \(\Delta\) are as defined in Lemma 4.1, is a solution for problem (12).

The next result states the properties of the ramp rate function for a tree network. The proof, given in the Appendix, is based on the description of the feasible region in terms of the power levels of the controllable nodes stated in Lemma A.1. For the ramp rate function with normally distributed loads defined in (12), one could obtain a similar result following Lemma 4.2 with \(\Delta\) replaced by \(\Delta'\) in the proof.

**Proposition 4.3.** (Ramp rate of tree network): Let \(G_m\) be a tree and \(H\) denote the hyperrectangle describing the region of operation of the controllable nodes, where opposite faces correspond to the minimum and maximum possible power level of a controllable node. Then the ramp rate \(R\) is piecewise affine on \(H\), i.e., for some \(s > 0\), \(H\) admits a decomposition

\[H = \{V_1 \cup V_2 \cup \ldots \cup V_s,\]

where \(\{V_s\}_{s=1}^n\) are polyhedra, and \(R\) is affine on each \(V_s\).

**Remark 1.** (Ramp rate for networks with non-overlapping loops): If the network is not a tree, then the flows corresponding to a power injection vector are not unique. Nevertheless,
the ramp rate for networks with non-overlapping loops is a non-increasing function of \( g \), as the feasible region of \( \|g\|_1 \) can only shrink with increase in some component(s) of \( g \).

Given a regulation power \( x \), we note that there may be more than one feasible operating point for the microgrid that produces it. As a result, the ramp rate as a function of regulation power is not uniquely defined. We address this by defining \( R : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0} \), as
\[
R(x) = \max_{g^*} \mathcal{R}(g^*),
\]
where \( g^* \) denotes a minimizer of the cost of producing the regulation \( x \) while respecting the power flow and capacity constraints. We take the maximum, since the optimizer \( g^* \) might not be unique. If the cost functions for all the controllable nodes are convex, each \( g^* \) is a decreasing function with respect to \( x \), which means that at least one component of \( g^* \) would decrease as \( x \) increases (using the convention that up regulation is negative). Using this fact, we conclude that \( R \) as a function of \( x \) is non-decreasing, with maximum possible value as \( 1^\top r \). Figure 3 provides the ramp rate functions of the four groups of microgrids displayed in Figure 3 in the constant load case.

![Ramp rate functions for different instantiations of the reduced-order UCSD microgrid with constant loads.](image)

Fig. 4. Ramp rate functions for different instantiations of the reduced-order UCSD microgrid with constant loads. The shaded regions represent the range of regulation power that the corresponding microgrid can provide.

In Remark 2, we discuss the conditions under which the minimum ramp rate of the microgrid is always non-zero.

**Remark 2. (Non-zero minimum ramp rate):** It is natural to argue that the microgrid could have a zero minimum ramp rate. Here, we discuss conditions under which the minimum ramp rate of the microgrid is non-zero. Let \( \mathcal{E}' = \{ e_j \in \mathcal{E} \mid \omega_j < \bar{\omega}_j \} \) be the set of all the lines which have not reached their flow limits when providing the maximum up regulation. Next, consider the graph \( G'_m = (\mathcal{V}, \mathcal{E}') \) and let \( V'_g = \{ v_i \in V_g \mid \exists \text{ a path from } i \text{ to } 1 \text{ in } G'_m \} \), i.e., the set of controllable nodes which are connected to the tie line. If \( V'_g \neq \emptyset \), then the minimum ramp rate is always non-zero.

### 4.3 Cost Function

Each aggregator needs to calculate the cost of providing a given amount of regulation by capturing the effect of operating the controllable nodes away from their baseline operating points. For an operating point \( g \), the total cost for the aggregator is given by
\[
h(g) = \sum_{p \in V_g} h_p(g_p),
\]
where \( h_p : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0} \) is the cost of operating node \( p \) away from its baseline level \( g^0_p \). One representative example of such a function is \( h_p(g_p) = (g_p - g^0_p)^2 \). The total regulation that the aggregator provides is the combination of individual regulations of controllable nodes. Therefore, for a specified regulation level \( x \), one would ideally choose the value of \( g \) that minimizes the total cost given by \( \text{(14)} \) respecting the power flow constraints in \( \text{(5)} \) and the minimum and maximum capacity constraints on each controllable node. Formally, \( f : [\pi, z] \times [\pi, z] \rightarrow \mathbb{R}_{\geq 0} \), is given by
\[
f(x) = \begin{cases}
\min_{g, \omega} & h(g) \\
\text{s.t.} & g \leq g \leq \bar{g} \\
& (P^0 + x) g^\top - l^\top = M \omega \\
& |\omega| \leq \bar{\omega}.
\end{cases}
\]

However, a cost function defined like this does not take into account the previous operating point of the microgrid and assumes that it can transition between the optimal points corresponding to different regulation powers arbitrarily fast. In practice, however, since the regulation set points change every 2-4 seconds, ramp rates might limit the change from optimal point at one time instant to the next. This suggests that the cost of providing certain amount of regulation at one instant also depends on the value of the regulation power at the previous instant. Hence, we define the cost \( f : [\pi, z] \times [\pi, z] \rightarrow \mathbb{R}_{\geq 0} \), of providing regulation power \( x \), if providing regulation power \( x^- \) at the previous instant, as
\[
\min_{g, \Delta g, \omega, \Delta \omega} h(g + \Delta g)
\]
\[
\text{s.t.} \quad g \leq g + \Delta g \leq \bar{g}, \quad \Delta g \leq r
\]
\[
\frac{[(P^0 + x) (g + \Delta g)^\top - l^\top]}{\omega + \Delta \omega} \leq \bar{\omega}
\]
\[
\frac{g \leq \bar{g}}{|\omega| \leq \bar{\omega}}
\]
\[
\frac{[(P^0 + x^-) g^\top - l^\top]}{M \omega} = M \omega.
\]

Here, \( (g + \Delta g, \omega, \Delta \omega) \) and \( (g, \omega) \) are the vectors of the power levels of controllable nodes and line flows when the microgrid provides regulation power \( x \) and \( x^- \), respectively. The constraints also enforce the capacity limits for the individual controllable nodes and the flow limit constraints for both values of regulation power, and the ramp constraints in transitioning from \( x^- \) to \( x \). The reason to include the power flow constraints at \( x^- \) in \( \text{(16)} \) is to enable the aggregator to pre-compute the cost function independently of the regulation power it might be asked to provide. Otherwise, if the cost is computed at every regulation instant, \( g \) and \( \omega \) providing \( x^- \) would be known, and the optimization variables would only be \( \Delta g \) and \( \Delta \omega \). As such, \( f(x, x^-) \) is a lower bound on the actual cost since \( (g, \omega) \) are also decision variables and are selected optimally to move to the next operating point.
The following result, whose proof is given in the Appendix, identifies a condition that simplifies the computation of the cost function $f(x, x^-)$ defined in (16).

**Lemma 4.4.** (Simplified formulation and convexity of cost function): Given regulation powers $x^-$ and $x$, if $|x - x^-| \leq R(x^-)$, then $f(x, x^-) = f(x)$. If $h$ is (strictly) convex, then $f$ is (strictly) convex.

Figure 5 provides the cost functions (15) of the four groups of microgrids displayed in Figure 3 in the constant load case.

![Cost Functions](image)

**TABLE 1**

| Bid Quantity | Value |
|--------------|-------|
| Capacity     | $\mathcal{F}$ |
| Mileage      | $k R'(g^{op})$ |
| Capacity price | $h(g^{op})/|\mathcal{F}|$ |

Note that the cost function (15) assumes the load to be constant, but since the aggregator is not required to submit its cost functions in [CP1], there is no need to pre-compute this using probabilistic techniques. Instead, the cost function at a given regulation instant could be computed online using the load at that instant. Time taken to compute the cost function at a given instant would highly depend upon the type of solver used but is usually very small (e.g. less than a second with built in MATLAB solver `fmincon`). It should also be noticed that although the cost function could be computed in a short span of time, but since the regulation period lasts for 10-15 minutes, in general, the variation in load would not be too much and hence, it would not be needed to compute it too often.

### 4.4 Bids for Participation in Market Clearance

Based on the abstractions in Sections 4.1-4.3 here we specify the bid information used by each aggregator to participate in [CP1]. Without loss of generality, we specify the bid quantities for up regulation market. Let $g^{op} \in \mathbb{R}^n_{>0}$ denote the component in $g$ of the solution of (2).

Table 1 specifies the proposed values for the bidding quantities. Here $k > 0$ is a constant depending on the duration of the regulation period and update frequency of the AGC setpoints. The suggested bids are conservative, meaning that the aggregator would be able to provide whatever it promises, and there is no strategy to maximize profit. It might seem from Table 1 that there is no need to compute beforehand the whole ramp rate function $R$ in Section 4.2. However, a risk taking aggregator might use a higher value of mileage bid based on the shape of $R$. It is also interesting to note that, from the convexity of cost function in Lemma 4.4 and the capacity price bid in Table 1 the aggregator would never be at loss regardless of the regulation power being provided.

### 5 RTO-DERP COORDINATION PROBLEM

Here we describe our algorithmic solution for the RTO-DERP coordination problem [P2] to disaggregate the regulation signal. Equipped with the microgrids’ capacities and cost and ramp rate functions identified in Section 4, the aggregators, communicating over a graph $\mathcal{G}$, seek to solve, at each instant of the regulation period, the optimization problem (2). However, as we have noted before, this problem might not always be feasible due to the presence of ramp constraints. This means that in principle, at each regulation instant, one would need to solve (2) if it is feasible or minimize the difference between the required regulation and the procured regulation if it is infeasible. Such dichotomy also raises the issue of the necessary information available to the aggregators to determine which one of the two cases to address at each regulation instant.

Instead, we propose to reformulate the optimization problem in a way that lends itself to the identification of solutions that minimize the error between the procured regulation and the required regulation whenever (2) is not feasible. Without loss of generality, throughout this section we assume the required regulation power to be positive. We start by defining the problem

$$\min_{x} \quad f^\mu(x) = f(x) + \mu|\Delta x|^+$$

s.t. $x_i \leq x_i \leq x_i^T \quad \forall i$

$$|x_i - x_i^-| \leq R_i(x_i^-) \quad \forall i,$$

where $\mu > 0$ is a penalty parameter and $\Delta x = x_r - x_i$. The following result, whose proof is given in the Appendix, characterizes the equivalence between problems (17) and (2).

**Lemma 5.1.** (Equivalence between (2) and (17)): Optimization (17) is always feasible and there exists $\bar{\mu} < \infty$ such that for all $\mu \in [\bar{\mu}, \infty)$, (2) and (17) have the same solution set if (2) is feasible.

**Remark 3.** (Establishing the threshold value $\bar{\mu}$ without the knowledge of dual optimizers): The threshold value $\bar{\mu}$ in Lemma 5.1 depends on the optimal values of the dual variables, which is not known beforehand. Interestingly, the explicit knowledge of the Lagrange multipliers to obtain a lower bound.
on the value of $\mu$ can be avoided. In fact, according to Proposition 5.2, we have

$$\hat{\mu} \geq 2 \max_{x \in F} \| \nabla f(x) \|_{\infty}.$$  

- Given Lemma 5.1 we focus on solving problem (17) in a distributed way. To handle the local constraints, we again reformulate (17) using exact penalty function as

$$\min_x f^P(x) = f(x) + \mu_2 \sum_{i=1}^{N} \big( \langle b_i, \Delta x \rangle + \mu [\Delta x]_+ \big), \quad (18)$$

where $b_i = x_i - \min (x_i, R_i (x_i)),$ and $b_i = \max (x_i - R_i (x_i), 0),$ are the box constraints taking care of the capacity and ramp rate for aggregator $i \in \{1, \ldots, N\}$ and $\mu_2 > 0$ is again a penalty parameter. Once again, similar to Lemma 5.1 there exist finite values of $\mu_2$ for which the reformulation (18) is exact.

Since problem (18) is unconstrained, consider the dynamics

$$\dot{x} \in -\partial f^P(x), \quad (19)$$

where $\partial f^P : \mathbb{R}^N \rightarrow \mathbb{R}^N$ denotes the generalized gradient of $f^P$. For each agent $i \in \{1, \ldots, N\}$, $\partial f^P(x)$, is given by

$$\nabla f_i(x_i) - [\mu_1 \Delta x - \mu_2]_{b_i} - [\mu_2]_{b_i} b_i - \mu_2 \mu_1 \Delta x \neq 0,$$

$$\nabla f_i(x_i) - [0, \mu] - [\mu_2]_{b_i} b_i - \mu_2 \mu_1 \Delta x \neq 0,$$

$$\nabla f_i(x_i) - [\mu_1 \Delta x - \mu_2]_{b_i} + [\mu_2]_{b_i} b_i - \mu_2 \mu_1 \Delta x \neq 0,$$

$$\nabla f_i(x_i) - [0, \mu] + [\mu_2]_{b_i} b_i - \mu_2 \mu_1 \Delta x \neq 0,$$

$$\nabla f_i(x_i) - [\mu_1 \Delta x - \mu_2]_{b_i} + [0, \mu_2] - \mu_2 \mu_1 \Delta x \neq 0,$$

$$\nabla f_i(x_i) - [0, \mu] - [\mu_2]_{b_i} b_i + [\mu_2]_{b_i} b_i - \mu_2 \mu_1 \Delta x \neq 0.$$ 

The equilibria of the dynamics (19) satisfy $0 \in -\partial f^P(x)$. Asymptotic convergence of (19) to the optimizers of (18) could be easily established using tools from non-smooth analysis, cf. [28] Proposition 14. However, the implementation of (19) requires every aggregator to have the knowledge of the total regulation at all times. To handle this, we use dynamic average consensus, cf. Section 2 to estimate the average of the difference between the required regulation and procured regulation from all the microgrids. Since $\frac{1}{N} \Delta x$ and $\Delta x$ have the same signs, we modify algorithm (19) using dynamic average consensus as follows

$$\dot{x} \in -\partial f^{\mu_2}(x) + [\mu]_{\Delta x}^{+}, \quad (20a)$$

$$\dot{\hat{z}} \in -\nu - \beta L z - v + \nu (x,e - x) + \partial f^{\mu_2}(x) - [\mu]_{\Delta x}^{+}, \quad (20b)$$

$$\dot{\hat{v}} = \nu \beta L z,$$  

where $L \in \mathbb{R}^{N \times N}$ is the Laplacian matrix of $G, z, v \in \mathbb{R}^N,$ $z_i$ is the ith aggregator’s estimate of $\frac{1}{N} \Delta x,$ $[\mu]_{\Delta x}^{+} \in \mathbb{R}^N$ with its ith element as $[\mu]_{\Delta x}^{+}$ and $e$ is the unit vector with only one entry as one and all others as zero. Note immediately that the algorithm (20) is distributed over the communication graph and only one aggregator needs to know the required regulation. We refer to (20) as “gradient descent + dynamic average consensus” algorithm, abbreviated as $\psi_{\text{gadac}}$. The equilibria for $x$ are the points satisfying $0 \in -\partial f^{\mu_2}(x) + [0, \mu]$. The next result, whose proof is given in the Appendix, characterizes the convergence properties of the $\psi_{\text{gadac}}$ algorithm.

**Theorem 5.2.** (Asymptotic convergence of the distributed dynamics to the optimizers): Let $G$ be strongly connected and weight-balanced, and the initial conditions satisfy $1^T v(0) = 0$ and $1^T z(0) - \Delta x(0) = 0,$ then there exists $\hat{\mu} < \infty$ such that the dynamics $\psi_{\text{gadac}}$ find the optimizers of (18) for all $\mu \in [\hat{\mu}, \infty).$

**Remark 4.** (Initialization of the distributed algorithm): For the dynamics $\psi_{\text{gadac}}$ to converge to the optimizers, Theorem 5.2 specifies requirements on the initial conditions. The requirement $1^T v(0) = 0$ could be implemented trivially by selecting $v(0) = 0.$ For the implementation of $1^T z(0) - \Delta x(0) = 0,$ the aggregators can simply choose $z(0) = 0$ and $x_i(0) = 0$ for all i, except for the aggregator who knows the required regulation for which $x_i(0) = x_r.$

### 6 Simulations

In this section, we provide our simulation results based on the abstractions of capacities, cost, and ramp rate developed in Section 4 and the RTO-DERP coordination algorithm (20) in Section 5. For the purpose of simulations, we consider a reduced-order model of the University of California, San Diego (UCSD) microgrid developed using the distributed feeder reduction algorithm in [29] and provided by the research group of Prof. Jan Kleissel. Compared to the full-order model of the UCSD microgrid [30] which is a radial, balanced network with 1289 buses (3869 nodes), the reduced-order model is a balanced tree network with 48 buses. The buses in the reduced-order model are obtained by retaining the key buses in the full order model which are the buses where the building loads aggregate or which have generators. Since the UCSD reduced-order model is balanced, we consider only one phase in our simulations. The reduced-order UCSD microgrid consists of 10 generators and 37 loads. Out of the 10 generators, 2 are gas turbines, 1 is a steam turbine and 7 are solar photovoltaic. The loads comprise of 34 building loads and 3 electric vehicle stations. We show the location of the buses on the geographical map of the campus in Figure 6. For constructing different microgrids, we use different baseline scenarios to instantiate the UCSD microgrid template. We consider 12 different microgrids, divided into 4 groups. The baseline values of generation and mean load is constant within a group. The 3 different scenarios within a group consist of (a) constant load, (b) variable load with failure probabilities $\epsilon' = 10^{-1}, \epsilon = 4.2 \times 10^{-5}$ and (c) variable load with failure probabilities $\epsilon' = 2 \times 10^{-1}, \epsilon = 8.4 \times 10^{-5}$. The abstracted regulation capacities and ramp rate functions of different microgrid groups are shown in Figures 3 and 4 resp. For cost functions, we consider quadratics for all the resources. The abstracted cost functions for different groups are shown in Figure 7.

We demonstrate the performance of the distributed algorithm (20) in two sets of simulations. To implement the continuous-time algorithm, we use a first-order Euler discretization with step size of 0.001 to show its practical feasibility. The values of $\mu, \mu_2, \beta$ and $\nu$ are taken to be 1000, 1100, 400 and 400, respectively. In the first simulation, cf. Figure 7 we consider one regulation instant and first
show the evolution of the proposed algorithm for required down regulation of 50000 kW, and compare it, for the same communication topology (undirected ring with few additional edges), against the (2-hop distributed) saddle-point dynamics [31] of the augmented Lagrangian for the equivalent reformulated problem as per [32]. As can be seen from the plots, the algorithm time required by the proposed algorithm to reach 1% band of the required regulation power in much less compared to the saddle-point dynamics. The time required does increase when the communication topology is changed to a directed ring - which is the worst possible topology for strongly connected graphs, but still remains less than a second, implying that the number of iterations is less than 1000.

For the second simulation, we consider the dynamic regulation test signal (RegD), available on the Pennsylvania-New Jersey-Maryland Interconnection (PJM) website [33]. Since the RegD signal on the PJM website is normalized and could be scaled as long as the problem remains feasible, we scale it by a factor of 50000 and then use our abstractions and clear the market according to [CP1]. Once the market is cleared, we use our algorithm to track the scaled RegD signal and compare it using the current algorithm of disaggregating the regulation signal described in [CP2]. For the sake of clarity, we show only the first 100 instants of the regulation period, and instead of contributions from each of the 12 microgrids, show the total contributions from the 4 groups. As we can see from Figure 8, when it is not possible to provide the required amount due to limits on ramp rates, both the proposed algorithm and the current algorithm try to provide as much regulation power as possible, and the tracking performance for both the algorithms is similar. But, if we compare the cost, the proposed algorithm with a cost of $9104 outperforms the current algorithm with a cost of $9638. This difference in cost comes from very different power contributions from the microgrids for the two algorithms. The proposed algorithm allocates the regulation signal to the microgrids based on their abstracted cost functions (cf. Figure 5), whereas current practice does not take them into account. It can be noticed...
in Figure 8 that, under current practice, if not capped by the cleared capacities, the power allocations for different microgrid groups have the same ratios for every regulation instant. For example, the shape of the regulation power curves for microgrid groups 1, 3 and 4 are similar and only differ in terms of scaling (by factors depending on the ratio of their procured mileages).

Fig. 7. Performance of the proposed distributed algorithm compared against the saddle-point dynamics for 50000 kW down regulation from 12 aggregators. The dashed lines for the proposed algorithm represent the algorithm evolution over a directed ring, and the black dashed lines represent 1% band of the required regulation power. Dynamics were implemented in discrete time with a step size of 0.001 and the values of $\mu$, $\mu_2$, $\beta$ and $\nu$ as 1000, 1100, 400 and 400, respectively.

Fig. 8. Performance of the proposed RTO-DERP coordination algorithm and the algorithm followed currently tested against first 100 updates of the PJM RegD signal. For microgrid groups, the solid lines represent the regulation power using the proposed algorithm and the dashed lines represent the regulation power using current practice. Although the tracking performance for both the algorithms is similar, contributions from individual microgrids differ substantially resulting into different costs.

7 Conclusions
We have considered the problem of providing frequency regulation services by aggregations of DERs. We have described the limitations of current practice and identified the challenges to overcome them with DER aggregators modeled as microgrids. We have developed meaningful abstractions for the capacity, cost of generation, and ramp rates by taking into account the power flow equations inside the microgrid. This provides enough information for the microgrids to participate in the market clearance stage. We have employed these abstractions to design a provably correct distributed algorithm that solves the RTO-DERP coordination problem to optimally disaggregate the regulation signal when the problem is feasible and minimize the difference between the required regulation and procured regulation when it is infeasible. Future work will extend our work to microgrids with more general topologies, incorporate AC power flow equations, and investigate smooth distributed algorithms to remove any chattering due to nonsmooth dynamics.

Appendix
Here we provide proofs of all the results stated in the paper.

Proof of Lemma 4.1
With the notation of the statement, (6) can be written as

$$|M_1(1^T l - 1^T g) + M_2 g - M_3 l + N_j \gamma| \leq \omega.$$  

Without loss of generality, let us for now consider only the following constraint in (6)

$$\Pr(|\xi_j^\circ| - \omega \leq 0) \geq 1 - \epsilon.$$  

where $\xi_j = (M_{1j}^T 1 - M_{3j}^T)l + (M_{2j} - M_{1j}^T)g + N_j \gamma$. Let $\xi_j^+ = \{\xi_j \in \mathbb{R} : \xi_j - \omega \leq 0\}$ and $\xi_j^- = \{\xi_j \in \mathbb{R} : -\xi_j - \omega \leq 0\}$. Then (21) is equivalent to

$$\Pr(\xi_j^+ \cap \xi_j^-) \geq 1 - \epsilon.$$  

We can further rewrite (22) as

$$\Pr(\xi_j^+ \cup \xi_j^-) \leq \epsilon \Rightarrow \Pr(\xi_j^+ \cup \xi_j^-) \leq \epsilon.$$  

We next break (23) down into single chance constraints. Using the fact that $\xi_j^+$ and $\xi_j^-$ are mutually exclusive, $\Pr(\xi_j^+ \cup \xi_j^-) = \Pr(\xi_j^+ \cap \xi_j^-)$. Therefore, (23) is equivalent to

$$\Pr(\xi_j^+ \cap \xi_j^-) \leq \epsilon/2, \text{ and } \Pr(\xi_j^- \cap \xi_j^+) \leq \epsilon/2.$$  

If $l \sim \mathcal{N}(\hat{l}, V_l)$, then $\xi_j \sim \mathcal{N}(\hat{\xi}_j, \sigma_j^2)$ where

$$\hat{\xi}_j = (M_{1j}^T 1 - M_{3j}^T)\hat{l} + (M_{2j} - M_{1j}^T)g + N_j \gamma,$$

$$\sigma_j^2 = (M_{1j}^T 1 - M_{3j}^T)V_l(M_{1j}^T 1 - M_{3j}^T)^\top.$$  

Defining $w = (\xi_j - \hat{\xi}_j)/\sigma_j$, we have $w \sim \mathcal{N}(0, 1)$. Then

$$\Pr(\xi_j^+) = \Pr\left(w \leq \frac{\omega - \hat{\xi}_j}{\sigma_j}\right) = \Phi\left(\frac{\omega - \hat{\xi}_j}{\sigma_j}\right).$$  

Using equations (25) and (1), we have from (24) for $\Pr(\xi_j^-)$

$$\frac{1}{2} + \frac{1}{2} \erf\left(\frac{\omega - \hat{\xi}_j}{\sqrt{2}\sigma_j}\right) \geq 1 - \epsilon/2,$$
\( \Rightarrow \text{erf} \left( \frac{\tilde{w}_j - \tilde{c}_j}{\sqrt{2} \sigma_j} \right) \geq 1 - \epsilon, \)
\( \Rightarrow \tilde{c}_j \leq \sqrt{2} \sigma_j \text{erf}^{-1}(\epsilon - 1) + \tilde{w}_j. \)

A similar inequality could be obtained from (24) for \( \text{Pr}(\xi^-) \).
As a result, (24) could be rewritten as
\( |\tilde{c}_j| \leq \sqrt{2} \sigma_j \text{erf}^{-1}(\epsilon - 1) + \tilde{w}_j. \)

The righthand side of the above constraint is a constant dependent on \( \epsilon \) and the left hand side depends on the decision variables \( g \) and \( \gamma \).

The same technique could be applied to the remaining set of constraints. If we apply this to all the chance constraints in (9), then problem (9) could be solved by solving the deterministic linear program (10). □

**Lemma A.1.** (Simplified power flow constraints for tree network): Let \( \mathcal{G}_m \) be a tree and \( P_{\text{ref}} \in \mathbb{R}^{(n-1) \times (n-1)} \) denote its path matrix with first vertex as reference ref. Then the constraints
\[
[(P - 1^T \Delta g) (g + \Delta g)^T - 1]^T = \mathcal{M}(\omega + \Delta \omega),
\]
\( |\omega + \Delta \omega| \leq \overline{\omega}, \)

in (11) could be equivalently written as
\[
P_1^T \Delta g \leq \overline{\omega} + P_2^T \Delta g + P_1^T g,
\]
where \( |P_1^T|, |P_2^T| = |P_{\text{ref}}^T| \), with \( P_1 \in \mathbb{R}^{n_g \times (n-1)} \) and \( P_2 \in \mathbb{R}^{n_l \times (n-1)} \), and \( |P_{\text{ref}}^T| \) denotes the non-negative matrix whose elements are given by the absolute values of the corresponding elements of \( P_{\text{ref}}^T \).

**Proof.** Let \( M_{\text{ref}} \in \mathbb{R}^{(n-1) \times (n-1)} \) denote the matrix obtained after removing the row corresponding to reference vertex from \( M \). According to (34), we have
\[
M_{\text{ref}}^{-1} = P_{\text{ref}}^T.
\]

With first vertex as ref, equation (53) could be rewritten as
\[
[g + \Delta g] = M_{\text{ref}}(\omega + \Delta \omega),
\]
where we have used the fact that rank(\( M \)) = rank(\( M_{\text{ref}} \)) = \( n - 1 \), cf. (24) Corollary 4-4]. Using (29) and (28), constraint (26) is equivalent to
\[
-\overline{\omega} \leq P_{\text{ref}}^T [g + \Delta g] \leq \overline{\omega}.
\]

Due to the structure of \( P_{\text{ref}}^T \), cf. Section 2 all the non-zero entries for any row of \( P_{\text{ref}}^T \) are either 1 or -1. Since we are characterizing the ramp up rate and are only concerned with what happens to the feasible region with the increase in some component(s) of \( g \), the active constraint for the lines for which the non-zero entries are 1 would be
\[
P_{\text{ref}}^T [g + \Delta g] \leq \overline{\omega},
\]
and for the lines for which the non-zero entries are -1 would be
\[
-P_{\text{ref}}^T [g + \Delta g] \leq \overline{\omega}.
\]

(30) is equivalent to (27), completing the proof. □

**Proof of Proposition 4.3**
Let us start by denoting the region where
\[
P_1^T r \leq \overline{\omega} + P_2^T \Delta g + P_1^T g,
\]
by \( V_1 \). Boundaries of \( V_1 \) are \( (n-1) \) hyperplanes given by
\[
P_1^T r = \overline{\omega} + P_2^T \Delta g + P_1^T g.
\]

Some of these hyperplanes could even be outside \( H \). But in general, all these \( (n-1) \) hyperplanes could be the faces of \( V_1 \). It is clear that in \( V_1 \), none of the flow constraints is active and \( R(g) = 1^T r \).

Outside \( V_1 \), we have
\[
\overline{\omega} + P_2^T \Delta g + P_1^T g < P_1^T r
\]
for at least one \( j \in \{1, \ldots, n-1\} \). First we consider the region where (31) holds for only one such \( j \), denoted as \( j' \). Then either
\[
\overline{\omega} + P_2^T \Delta g + P_1^T g \geq 0, \text{ or } \overline{\omega} + P_2^T \Delta g + P_1^T g = 0.
\]
In the former case, we are in the polyhedron whose two faces are given by
\[
\overline{\omega} + P_2^T \Delta g + P_1^T g = 0, \text{ and } \overline{\omega} + P_2^T \Delta g + P_1^T g = 0.
\]
Let us call one of these polyhedron \( V_2 \). In \( V_2 \), \( R(g) = 1^T \Delta g \), where \( \Delta g \) satisfies
\[
\overline{\omega} + P_2^T \Delta g + P_1^T g = P_{1j}^T \Delta g.
\]
For \( 1^T \Delta g \) to be maximum, the controllable nodes for which the corresponding entries are zero in \( P_{1j} \), we will have \( \Delta g_{pj} = r_p \). As some component(s) of \( g \) for which the corresponding entry in \( P_{1j} \) is 1 increases, some components of \( \Delta g \) with corresponding entry 1, decrease to balance it. Hence, \( R(g) = 1^T r - P_{1j'}^T g \).

Now considering the latter case when
\[
\overline{\omega} + P_2^T \Delta g + P_1^T g = 0.
\]

On this hyperplane, \( R(g) \) becomes constant again as the controllable nodes for which the corresponding entries are zero in \( P_{1j'} \) have \( \Delta g_{pj} = r_p \) and other entries of \( \Delta g \) have to be zero. Hence, \( R(g) = (1 - P_{1j'})^T r \). Note that different polyhedrons similar to \( V_2 \) might exist with different \( j' \).

Now we consider the regions where (31) holds for multiple \( j \in \{1, \ldots, n-1\} \). Let us denote by \( V_3 \) the polyhedron, whose few faces are given by
\[
\overline{\omega} + P_2^T \Delta g + P_1^T g = P_{1j}^T r,
\]
for all \( j \) satisfying (31). Inside \( V_3 \), \( R(g) = 1^T \Delta g \), where \( \Delta g \) is given by the simultaneous solution of
\[
P_{1j}^T \Delta g \leq \overline{\omega} + P_2^T \Delta g + P_1^T g,
\]
for all the corresponding \( j \) and \( 1^T \Delta g \) is maximum. At least, one of these inequalities would hold with equality. Similar to \( V_2 \), we notice that if we increase some component(s) of \( g \) in \( V_3 \) with corresponding entry in any of \( P_{1j} \) as 1, \( R(g) \) decreases linearly. While increasing some component of \( g \), a point would be reached where
\[
\overline{\omega} + P_2^T \Delta g + P_1^T g = 0.
\]
for some $j$ and that would be another face of $V_3$. On this hyperplane, $\Delta g_\rho = 0$ for the controllable nodes for which the corresponding entry of $P_{1j} = 1$ in $[52]$. Note that $\mathcal{R}(g)$ is still linear as $V_3$ but with a different slope.

In general, depending on the parameters of the microgrid at hand, there would be several polyhedrons where $[31]$ holds for different $j$. But the characterization of ramping capacity would be similar to $V_3$ in all these polyhedra. Since, the ramp rate is either affine or constant in all the polyhedra, it is affine. 

Proof of Lemma 4.4

If the difference between two regulation powers, i.e., $|x-x^-$ is greater than the ramp rate at $x^-$, then the microgrid might not be able to provide the regulation power at all. On the other hand, if the difference is less than the ramp rate, then it is clear that the microgrid would be able to provide the required regulation power optimally. So, in the latter case, the cost of providing regulation power $x$ or the solution of $[15]$ is equivalent to the optimization in $[15]$.

Next, we provide a proof for the convexity of $f$ if $h$ is convex. Let $C(x) = C_0 \cap C_1(x)$, where $C_0$ denotes the capacity constraints for $g$ and $C_1(x) = \{g \mid \begin{bmatrix} \rho_0 + x \\ g \\ -l \end{bmatrix} = M \omega \text{ and } |\omega| \leq \omega \}$.

Then, we have

$$f(x) = \min_{g \in C(x)} h(g).$$

Let $x_1, x_2 \in [\bar{x}, \bar{z}]$, where $\bar{x}$ and $\bar{z}$ are respectively, the maximum up and down regulation identified in Section 4.1. Then $f(x_1) = \min_{g \in C(x_1)} h(g)$, which means that for all $\delta > 0$, there exists $g_1 \in C(x_1)$ such that $f(x_1) + \delta \geq h(g_1)$. Similarly, there exists $g_2 \in C(x_2)$ such that $f(x_2) + \delta \geq h(g_2)$. Since $g_1 \in C(x_1)$ and $g_2 \in C(x_2)$, therefore $\lambda g_1 + (1 - \lambda)g_2 \in C(\lambda x_1 + (1 - \lambda)x_2)$, where $\lambda \in [0, 1]$. Hence,

$$f(\lambda x_1 + (1 - \lambda)x_2) = \min_{g \in C(\lambda x_1 + (1 - \lambda)x_2)} h(g),$$

$$\leq h(\lambda g_1 + (1 - \lambda)g_2),$$

$$\leq \lambda h(g_1) + (1 - \lambda)h(g_2),$$

$$\leq \lambda f(x_1) + (1 - \lambda)f(x_2) + \delta,$$

where the second last inequality would be strict in case of strict convexity. Since $\delta$ is arbitrary, it follows that $f$ is (strictly) convex. 

Proof of Lemma 5.4

We begin by noting that $x_i = x_i^-$ for each $i$ satisfies both set of constraints in $[17]$, since $x^-$ is the set of regulations provided by the aggregators at the previous instant. Hence, $[17]$ is always feasible. To prove the equivalence between the two problems, as our first step, we rewrite $[2]$ as

$$\begin{align*}
\min_x & \quad f(x) \\
\text{s.t.} & \quad x_r \leq 1^T x, \\
& \quad x_i \leq \bar{x}_i \quad \forall i, \\
& \quad |x_i - x_i^-| \leq R_i(x_i^-) \quad \forall i.
\end{align*}$$

Note that the equality constraint in $[2]$ is replaced by the inequality constraint in $[33]$.

For simplicity of exposition, we ignore the box constraints and write $[20]$ as

$$\begin{align*}
\dot{x} &= -\nabla f(x) + [\mu]^+ \\
\dot{z} &= -\nu z - \beta L z - \nu (x_{r,e} - x) + \nabla f(x) - [\mu]^+, \\
\dot{v} &= \nu \beta L z,
\end{align*}$$

First, consider the function $V_2(x, z) = 1^T z - \Delta x$. The derivative $L_{\psi, \Delta} V_2 : \mathbb{R}^{2N} \rightarrow \mathbb{R}$ is then given by

$$L_{\psi, \Delta} V_2 = 1^T \dot{x} + 1^T \dot{z} = -\nu 1^T (z - (x_{r,e} - x)) = -\nu V_2,$$

where we have used the fact that $1^T v = 0$ due to the initial condition $1^T v(0) = 0$ and dynamics $[34]$. The above equation implies that the summation of all the entries of $z$ converges to the mismatch between the required regulation and procured regulation exponentially with rate $\nu$. Hence $1^T z - \Delta x \equiv 0$ with the stated initialization.

Next consider the change of coordinates $(x, z, v) \rightarrow (x, z, \eta)$, with $\eta = \nu (z - (x_{r,e} - x)) + v$. The dynamics for $z$ and $\eta$ are then given by

$$\begin{align*}
\dot{x} &= -\beta L z - \eta + \nabla f(x) - [\mu]^+, \\
\dot{\eta} &= -\nu \eta.
\end{align*}$$

Consider the Lyapunov function candidate $V : \mathbb{R}^{3N} \rightarrow \mathbb{R}_{\geq 0}$,

$$V(x, z, \eta) = f^\pi(x) + \sum_{i=1}^N \left( z_i + \frac{1}{2} \|\eta\|^2 \right),$$

whose generalized gradient $\partial V : \mathbb{R}^{3N} \rightarrow \mathbb{R}^{3N}$ is given by

$$\begin{align*}
\partial V(x, z, \eta) = & \left( \begin{array}c \\
\nabla f(x) - [\mu]^+ \\
\nabla f(x) - [0, \mu]^+ \\
\nabla f(x) - [0, 0, \mu]^+ \\
\nabla f(x) - [0, 0, 0, \mu]^+
\end{array} \right), \\
\Delta x \neq 0, z \neq 0, \\
\Delta x = 0, z \neq 0, \\
\Delta x = 0, z = 0.
\end{align*}$$

Following $[28]$, set-valued Lie derivative $L_{\psi, \Delta} V : \mathbb{R}^{3N} \rightarrow \mathbb{R}$ can then be computed as

$$L_{\psi, \Delta} V(x, z, \eta) = \begin{cases} 
-(\nabla f - [\mu]^+)^T \nabla f + [\mu]^+ \frac{2}{\nu \|\eta\|^2}, \\
\nu \|\eta\|^2, \\
\phi,
\end{cases}$$

otherwise.
We now analyze various cases of $\Delta x \neq 0, z \neq 0$ in the following:

**Case 1:** $\Delta x < 0$ and $z < 0$.

$$L^{\psi_{\text{glac}}} = -\|\nabla f\|^2 - v\|\eta\|^2.$$

**Case 2:** $\Delta x > 0$ and $z > 0$.

$$L^{\psi_{\text{glac}}} = -\|\nabla f\|^2 + 3\mu \nabla f^\top 1 - 2N\mu^2 - \eta^\top 1 - \nu\|\eta\|^2.$$

**Case 3:** $\Delta x > 0$ and $z \neq 0$.

$$L^{\psi_{\text{glac}}} = -\|\nabla f\|^2 - 2N\mu^2 + \nabla f^\top (\mu 1 + 2\mu z) - \beta([\mu z]_+) L z - \eta^\top [\mu z]_+ - \nu\|\eta\|^2,$$

where $N_p$ is the number of positive elements of $z$.

**Case 4:** $\Delta x < 0$ and $z \neq 0$.

$$L^{\psi_{\text{glac}}} = -\|\nabla f\|^2 + 2\nabla f^\top [\mu z]_+ - \beta([\mu z]_+) L z - \eta^\top [\mu z]_+ - N_p\mu^2 - \nu\|\eta\|^2.$$

We do not need to consider the case when $\Delta x > 0$ and $z < 0$ since $1^\top z - \Delta x \equiv 0$ due to the discussion above. Out of the 4 cases, $L^{\psi_{\text{glac}}} < 0$ for Case 1. For the remaining cases, since $f$ is globally proper and $\|\nabla f\|$ is bounded over any compact set, $L^{\psi_{\text{glac}}} < 0$ if the value of $\mu$ is taken large enough for the worst-case scenario ($N_p = 1$). Since $\max \phi = -\infty, \max L^{\psi_{\text{glac}}} > 0$ except at the equilibrium. This along with the fact that $V$ is locally Lipschitz and regular implies that $V$ satisfies the hypothesis of [28] Theorem 1. Hence, the dynamics $\psi_{\text{glac}}$ converge to the optimal solution asymptotically.

**REFERENCES**

[1] P. Srivastava, C.-Y. Chang, and J. Cortés, “Participation of microgrids in frequency regulation markets,” in American Control Conference, Milwaukee, WI, May 2018, pp. 3834–3839.

[2] CAISO, “Expanded metering and telemetry options phase 2 - distributed energy resource provider,” 2015, draft proposal electronically available at https://www.caiso.com/Documents/DraftFinalProposal_ExpandedMetering_telemetryOptions/hasbe_DistributedEnergyResourceProvider.pdf.

[3] “Order No. 2222: Participation of distributed energy resource aggregations in markets operated by regional transmission organizations and independent system operators,” Sep. 2020, available at https://www.ferc.gov/sites/default/files/2020-09/E1_0.pdf.

[4] “Order No. 725: Frequency regulation compensation in the organized wholesale power markets,” 2011, available at http://www.ferc.gov/whats-new/comm-meet/2011/102011/E-28.pdf.

[5] M. Kintner-Meyer, “Regulatory policy and markets for energy storage in North America,” Proceedings of the IEEE, vol. 102, no. 7, pp. 1065–1072, 2014.

[6] J. L. Mathieu, S. Koch, and D. S. Callaway, “State estimation and control of electric loads to manage real-time energy imbalance,” IEEE Transactions on Power Systems, vol. 28, no. 1, pp. 430–440, 2013.

[7] P. Codani, M. Petit, and Y. Perez, “Missing money for EVs: Economics impacts of TSO market designs,” available at https://esrm.com/abstracts/2529280.

[8] B. M. Sanandaji, H. Hao, K. Poola, and T. L. Vincent, “Improved battery models of an aggregation of thermally controlled loads for frequency regulation,” in American Control Conference, Portland, OR, 2014, pp. 38–45.

[9] J. T. Hughes, A. D. Domínguez-García, and K. Poola, “Identification of virtual battery models for flexible loads,” IEEE Transactions on Power Systems, vol. 31, no. 6, pp. 4660–4669, Nov 2016.

[10] S. Rahnama, J. Stoustrup, and H. Rasmussen, “Integration of heterogeneous industrial consumers to provide regulating power to the smart grid,” in IEEE Conf. on Decision and Control, Florence, Italy, 2013, pp. 6268–6273.

[11] O. Borne, M. Petit, and Y. Perez, “Provision of frequency-regulation reserves by distributed energy resources: Best practices and barriers to entry,” in International Conference on the European Energy Market (EEM), June 2016, pp. 1–7.

[12] E. Dall’Anese, S. Gugglia, A. Simonetto, Y. C. Chen, and S. V. Dhople, “Optimal regulation of virtual power plants,” IEEE Transactions on Power Systems, vol. 33, no. 2, pp. 1868–1881, 2018.

[13] B. Biegel, P. Andersen, T. S. Pedersen, K. M. Nielsen, J. Stoustrup, and L. H. Hansen, “Smart grid dispatch strategy for on/off demand-side devices,” in European Control Conference, Zürich, Switzerland, 2013, pp. 2541–2548.

[14] J. T. Hughes, A. D. Domínguez-García, and K. Poola, “Coordinating heterogeneous distributed energy resources for provision of frequency regulation services,” in Hawaii International Conference on System Sciences, Big Island, HI, January 2017, pp. 2983–2992.

[15] C.-Y. Chang, S. Martínez, and J. Cortés, “Grid-connected microgrid participation in frequency-regulation markets via hierarchical coordination,” in IEEE Conf. on Decision and Control, Melbourne, Australia, Dec. 2017, pp. 3501–3506.

[16] X. Su, S. C. Utoni, A. D. Domínguez-García, and P. W. Sauer, “Coordination of distributed energy resources in lossy networks for providing frequency regulation,” in IREP Bulk Power System Dynamics and Control Symposium, Espinho, Portugal, August 2017.

[17] R. Ghaemi, M. Abbaszadeh, and P. Bonanni, “Scalable optimal flexibility control of distributed loads in the power grid,” in American Control Conference, Milwaukee, WI, June 2018, pp. 6646–6651.

[18] P. MacDougall, A. M. Kosek, H. Bindner, and G. Deconinck, “Applying machine learning techniques for forecasting flexibility of virtual power plants,” in IEEE Electrical Power and Energy Conference (EPEC), Ottawa, ON, Canada, Oct 2016, pp. 1–6.

[19] Y. Wang, X. Ai, Z. Tan, L. Yan, and S. Liu, “Interactive dispatch modes and bidding strategy of multiple virtual power plants based on demand response and game theory,” IEEE Transactions on Smart Grid, vol. 7, no. 1, pp. 510–519, Jan 2016.

[20] T. Anderson, M. Muralidharan, P. Srivastava, H. V. Haghif, J. Cortés, J. Kleissl, S. Martínez, and B. Washom, “Frequency regulation with heterogeneous energy resources: A realization using distributed control,” IEEE Transactions on Smart Grid, 2020, submitted.

[21] S. S. Kia, J. Cortés, and S. Martínez, “Dynamic average consensus under limited control authority and privacy requirements,” International Journal on Robust and Nonlinear Control, vol. 25, no. 13, pp. 1941–1966, 2015.

[22] J. Bushnell, S. M. Harvey, and B. F. Hobbs, “Opinion on pay-for-performance regulation,” Market Surveillance Committee, California ISO, Tech. Rep., March 9 2012. [Online]. Available: http://www.ferc.gov/ Documents/Oppinion-Pay-for-Performance.pdf.

[23] O. Ajala, M. Almeida, I. Celanovic, P. R. Sauer, and A. D. Domínguez-García, “A hierarchy of models for microgrids with grid-fed feeding inverters,” in IREP Bulk Power System Dynamics and Control Symposium, Espinho, Portugal, August 2017.

[24] Q.-C. Zhong and T. Hornik, Control of power inverters in renewable energy and smart grid integration. John Wiley & Sons, 2012, vol. 97.

[25] M. Zhlobarysov and A. D. Domínguez-García, “Convex relaxations of the network flow problem under cycle constraints,” IEEE Transactions on Control of Network Systems, vol. 7, no. 1, pp. 64–73, 2020.

[26] S. Seshu and M. B. Reed, Linear Graphs and Electrical Networks. Addison-Wesley Publishing Company, 1961.

[27] A. Cherukuri and J. Cortés, “Distributed generator coordination for initialization and anytime optimization in economic dispatch,” IEEE Transactions on Control of Network Systems, vol. 2, no. 3, pp. 226–237, 2015.

[28] J. Cortés, “Discontinuous dynamical systems – a tutorial on solutions, nonsmooth analysis, and stability,” IEEE Control Systems, vol. 28, no. 3, pp. 36–73, 2008.

[29] Z. K. Pecenak, V. R. Disfani, M. J. Reno, and J. Kleissl, “Multiphase distribution feeder reduction,” IEEE Transactions on Power Systems, vol. 29, no. 2, pp. 1320–1328, March 2014.

[30] B. Washom, J. Dilliot, D. Weil, J. Kleissl, N. Balac, W. Torre, and C. Richter, “Ivory tower of power: Microgrid implementation at the University of California, San Diego,” IEEE Power and Energy Magazine, vol. 11, no. 4, pp. 28–32, 2013.

[31] A. Cherukuri, B. Gharesifard, and J. Cortés, “Saddle-point dynam-
ics: conditions for asymptotic stability of saddle points,” *SIAM Journal on Control and Optimization*, vol. 55, no. 1, pp. 486–511, 2017.

[32] A. Cherukuri and J. Cortés, “Distributed algorithms for convex network optimization under non-sparse equality constraints,” in *Allerton Conf. on Communications, Control and Computing*, Monticello, IL, Sep. 2016, pp. 452–459.

[33] PJM, “Dynamic regulation test signal (RegD) signal,” [http://www.pjm.com/~media/markets-ops/ancillary/regn-test-wave.ashx](http://www.pjm.com/~media/markets-ops/ancillary/regn-test-wave.ashx).

[34] J. Resh, “The inverse of a nonsingular submatrix of an incidence matrix,” *IEEE Transactions on Circuit Theory*, vol. 10, no. 1, pp. 131–132, 1963.

[35] D. P. Bertsekas, “ Necessary and sufficient conditions for a penalty method to be exact,” *Mathematical Programming*, vol. 9, no. 1, pp. 87–99, 1975.

**Priyank Srivastava** received the B.Tech degree in electrical engineering from National Institute of Technology, Kurukshetra, India in 2012, and the M.Tech degree in control & automation from Indian Institute of Technology Delhi, India in 2016. He is currently a Ph.D. candidate in mechanical engineering at the University of California, San Diego, USA. His current research interests include dynamical systems, distributed and fast algorithms for constrained network problems, and coordination of distributed energy resources to enable their participation in energy markets.

**Chin-Yao Chang** received the B.S. degree in Physics from National Taiwan University, Taipei, Taiwan, R.O.C. He received the Ph.D. degree in Electrical and Computer Engineering at the Ohio State University, Columbus, OH, USA, in July 2016. He was a postdoctoral researcher with the Mechanical and Aerospace Engineering at the University of California, San Diego, CA, USA, from 2016 to 2018. He is now a research engineer at National Renewable Energy Laboratory (NREL), Golden, CO, USA. His research interests include demand response, distributed optimization, cyber physical systems, optimal power flow, microgrid control, and quantum computing.

**Jorge Cortés** (M’02, SM’06, F’14) received the Licenciatura degree in mathematics from Universidad de Zaragoza, Zaragoza, Spain, in 1997, and the Ph.D. degree in engineering mathematics from Universidad Carlos III de Madrid, Madrid, Spain, in 2001. He held postdoctoral positions with the University of Twente, Twente, The Netherlands, and the University of Illinois at Urbana-Champaign, Urbana, IL, USA. He was an Assistant Professor with the Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA, USA, from 2004 to 2007. He is currently a Professor in the Department of Mechanical and Aerospace Engineering, University of California, San Diego, CA, USA. He is the author of Geometric, Control and Numerical Aspects of Nonholonomic Systems (Springer-Verlag, 2002) and co-author (together with F. Bullo and S. Martínez) of Distributed Control of Robotic Networks (Princeton University Press, 2009). He is a Fellow of IEEE and SIAM. At the IEEE Control Systems Society, he has been a Distinguished Lecturer (2010–2014), and is currently its Director of Operations and an elected member (2018–2020) of its Board of Governors. His current research interests include distributed control and optimization, network science, resource-aware control, nonsmooth analysis, distributed decision making, network neuroscience, and multi-agent coordination in robotic, power, and transportation networks.