

Distribution Feeder-Scale Fast Frequency Response via Optimal Coordination of Net-load Resources—Part I: Solution Design

Blake Lundstrom, *Senior Member, IEEE*, Sourav Patel, *Student Member, IEEE*, and Murti V. Salapaka, *Fellow, IEEE*

Abstract—This paper is the first of a two-part series that develops and experimentally demonstrates a first-of-its-kind hierarchical control solution for optimally dispatching thousands of deferrable loads and distributed energy resources (DERs) across a distribution feeder to provide fast frequency response (FFR) within 500 ms to the bulk power system. This approach rapidly coordinates resources online after a frequency event occurs, allowing fast-changing, behind-the-meter (BTM) resources to be incorporated and aggregate FFR power set points to be achieved more quickly and accurately than existing approaches. We also present a solution for determining the optimal amount of headroom to operate solar inverters with to minimize opportunity cost while ensuring the FFR response viability of a building with the inverter and deferrable loads. In Part I, we develop practical algorithms for fast, cost-based optimal dispatch at multiple aggregation scales (single building, multiple buildings, and full distribution feeder), establish their optimality, and demonstrate via simulation that they are faster than state-of-the-art, coordinated frequency response approaches. In Part II, the entire platform is implemented and experimentally verified using a unique power hardware-in-the-loop demonstration, including more than 100 powered loads and DERs connected to a real-world distribution network model and over 10,000 net-load resources dispatched.

Index Terms—fast frequency response (FFR), frequency control, optimal dispatch, flexible load, distributed energy resources (DER), demand response, coordinated response

NOMENCLATURE

System Variables

Total active power consumption	P_{Gg}
FFR response command	$\Delta P_{Gg}^* = P_{Gg}(t) - P_{Gg}(t_0)$
Total net-load groups (NLGs)	$k = 1 \dots N_{nlg}$
Net-load units per NLG _k	$j = 1 \dots N_{u,k}$
Deferrable loads per NLU _{k,j}	$i = 1 \dots N_{Ld,k,j}$

B. Lundstrom (e-mail: blake.lundstrom@nrel.gov) is with the Power Systems Engineering Center, National Renewable Energy Laboratory, Golden, CO, USA and the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN, USA.

S. Patel and M.V. Salapaka are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN, USA.

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the Advanced Research Projects Agency-Energy (ARPA-E) under grant no. DE-AR0000701. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

Net-load Unit Variables

	Single NLU	General multi-building system
Total nondeferrable load	P_L	$P_{L,k,j}$
Deferrable loads		
count	N_{Ld}	$N_{Ld,k,j}$
load power	$\hat{P}_{Ld,i} = (1 - x_i)P_{Ld,i}$	$\hat{P}_{Ld,kji}$
on/off	$x_i = \begin{cases} 1 & \text{off (deferred)} \\ 0 & \text{on} \end{cases}$	x_{kji}
participation unit cost	c_i	c_{kji}
Inverter		
pre-event power	$P_{inv}(t_0) = P_{inv,base}$	$P_{inv,k,j}(t_0)$
post-dispatch power	$P_{inv} = P_{inv,base} + P_{inv\Delta}$	$P_{inv,k,j,base}$
limited change in power	$P_{inv\Delta,dn} \leq P_{inv\Delta} \leq P_{inv\Delta,up}$	$P_{inv\Delta,k,j}$
downward headroom	$P_{inv\Delta,dn}$	$P_{inv\Delta,k,j,dn}$
upward headroom	$P_{inv\Delta,up}$	$P_{inv\Delta,k,j,up}$
max power	P_{inv}^{max}	$P_{inv,k,j}^{max}$
participation unit cost	c_{inv}	$c_{inv,k,j}$
Grid power consumption	P_G	
FFR response command	$\Delta P_G^* = P_G(t) - P_G(t_0)$	

I. INTRODUCTION

A. Motivation and Background

RELIABLE operation of a power system depends on maintaining its frequency close to its nominal value at all times. In many grids worldwide, the displacement of conventional, high-inertia, synchronous generation by inverter-interfaced low-zero-inertia renewable generation has resulted in a decrease in total system inertia and a decline in frequency response performance from historic values. These trends are especially prevalent in isolated grids with high penetrations of renewable generation, such as Ireland/Northern Ireland and South Australia [1], but also evident in much larger interconnections in Europe [2] and the United States [3]. The outcome is that sudden imbalances in grid load and generation result in faster rates of change of frequency and larger frequency nadirs, making frequency regulation more challenging.

With continued grid modernization, new approaches to ensure sufficient frequency regulation are required. Traditionally, mitigation of frequency deviations after a sudden loss of generation or load is driven by three primary factors: inertial response, primary frequency response (PFR), and secondary frequency response (SFR). System operators, such as ERCOT in Texas, USA and ENTSO-E in Europe, have conducted recent studies (e.g., [2], [4]) to examine the decreasing inertia in their systems and propose mitigation measures. ERCOT's study highlights faster responding reserves as a key mitigation measure. Frameworks in different regions vary but generally include Fast Frequency Response (FFR), a response of the resource's full allocated capacity within 1 s or less after the

frequency crosses a predefined threshold or a dispatch signal is received. FFR is particularly effective because it can increase the time to system frequency nadir, allow additional time for PFR to deploy, and contribute to reducing the nadir's magnitude [5], [6]. In [5], a study on the ERCOT system found that FFR can be up to 2.2 times more effective than PFR, depending on the system inertia. As of 2019, many system operators have implemented or are finalizing the implementation of new FFR ancillary service market products that are distinct from PFR (e.g., FFR under ERCOT's NPPRR863 or the U.K. National Grid's Enhanced Frequency Response Service).

Demand-side resources, including DER and deferrable load ("net-load" resources), offer a largely untapped opportunity to provide valuable FFR services. It has been shown that system frequency disturbances propagate through distribution systems much faster than transmission systems [7]; thus, a collection of devices connected to a distribution system will observe, and can respond to, frequency disturbances just as fast (if not faster) than a traditional frequency response asset (e.g., sync. generator) connected at the same location would. DERs, such as solar inverters, have traditionally not contributed to frequency response [3], but recent work demonstrated new control methods to allow BTM solar inverters [8] and battery inverters [9] to rapidly respond (within 75 ms) to changes in grid frequency. At the solar power plant scale, the provision of FFR from multiple +20-MW solar power plants was demonstrated in [6]. Note that providing FFR from a solar inverter represents an opportunity cost; to provide increased active power output in response to a system loss of generation, the inverter must nominally operate off of its maximum power point (leaving "headroom" to increase the power output), forgoing some energy production and the associated revenue. Determining the optimal amount of headroom to leave for a given solar inverter's operating conditions is not studied in the literature. Here, we propose an approach to determine the optimal headroom. Deferrable load is particularly well suited to provide FFR services because it does not incur an opportunity cost in preparing for FFR provision, and it can respond with its full capacity as quickly as a relay can be operated or a modern appliance's power electronic front-end can be controlled. In particular, thermostatically controlled loads (TCLs) and electric vehicles (EVs) are well-suited because of their inherent energy storage properties, which allow them to be switched off for a short period (typically at least a few minutes) without significantly compromising user comfort [10], [11]. TCLs represent a significant portion of total system load: more than 48% of residential electricity consumption in the U.S. in 2018 [12].

A key challenge to leveraging demand-side resources to provide FFR in associated ancillary service markets is the ability to provide sufficient capacity (historically, ≥ 1 -MW capacity is required) and reliable dispatchability to a particular set point. Recently, regulators are considering actions to remove barriers to the participation of DER aggregations in ancillary service markets, including reducing minimum capacity requirements [13]. Coordinating deferrable load and DERs as a virtual power plant (VPP) capable of FFR is particularly attractive because it not only allows aggregations

of resources to meet market capacity participation limits but also leverages diversity, manages the variability of load and generation to enable dispatchability, and provides a framework to incentivize BTM participation and achieve the lowest-cost global dispatch solution. Also, by coordinating variable DERs and load in the same framework, a better determination of optimal DER operating headroom can be made.

B. Related Work

Few existing works address the coordination of net-load resources to provide FFR given the relatively new definition of FFR as a grid ancillary service product. For context, we examine related work, including existing PFR response approaches, which are often uncoordinated, as well as more general frequency response approaches that are coordinated.

Using deferrable load to contribute to grid services such as peak demand reduction and load following was explored as early as the late 1970s, though work on controlling loads fast enough to provide PFR has occurred only more recently. Earlier work includes a simulation study [14] that implemented autonomous PFR via temperature-based frequency trip points at individual refrigerators; simulation results demonstrated how autonomous PFR from 1,000 refrigerators could, in aggregate, provide PFR similar to that of a synchronous generator to the bulk power system. In [15], more than 50 electric water heaters and 150 clothes dryers were controlled in one of the earlier field demonstrations of fast (~ 400 -ms response times), autonomous PFR from loads. Later simulation studies identified overresponse, frequency oscillation, and loss of load diversity as potential issues with uncoordinated load frequency response [16]–[19]. Strategies to mitigate potential issues in autonomous approaches have been proposed, including using uniformly distributed response time delays [11] and implementing random activation and re-activation frequency thresholds [20]. These *uncoordinated* PFR approaches demonstrate meaningful contribution to system PFR, but they are unable to provide dispatchability to the operator and will result in uncertainty of response and suboptimal dispatch.

Semicoordinated load PFR approaches (e.g., [17], [21]–[24]), in which individual resources respond rapidly via autonomous control while having their control parameters periodically updated for coordination purposes, provide some improvement. Reference [17] proposed a supervised threshold determination approach in which electric water heaters are given updated frequency thresholds every 15 minutes and a given aggregate droop characteristic is achieved. Reference [24] demonstrated a similar approach applied to HVAC equipment with room temperature as feedback and updates made every minute. Reference [23] developed an hourly optimization-based approach that accounts for individual resource costs, comfort constraints, and global cost. Reference [21] extended the approach in [23] by adding load priorities and appliance off-time constraints to the optimization problem. Reference [22] solved a minimum cost commitment problem to compute the lowest cost interruptible load action to take for every given frequency deviation with hourly updates. These semicoordinated PFR approaches add increased capabilities to manage the penetration level and specific settings of resources participating autonomously to achieve a more predictable overall

PFR. However, they all assume simplistic, slow-changing, aggregated load models and perform coordination updates on a minute to hour basis, resulting in significant uncertainty in their dispatch when applied to controlling real-world net-load resources that change their active power significantly faster. Given this uncertainty, it is unlikely that such approaches will be able to accurately achieve an aggregate power set point and are much more likely to make sub-optimal dispatch decisions.

Online coordination approaches, in which individual resources are coordinated after a frequency event occurs, often have the advantage of optimal dispatch and the ability to integrate with automatic generation control (AGC), but they are generally too slow to meet PFR or FFR timescales. For example, in [25], eight optimization-based coordination frameworks in the literature are compared, and all have computation times of 10 s or greater. Distributed approaches are used to solve an optimal load-shedding problem [26] and an optimal load-control problem [27] on the +5-s timescale. Reference [28] developed a distributed approach for coordinating load aggregators that integrates with AGC for SFR. Reference [29] developed a centrally-managed solution where aggregated loads are controlled autonomously using control parameters pre-coordinated based on day-ahead estimates to provide PFR and a SFR supervisory control can be integrated with AGC. Reference [30] demonstrated similar combined SFR and PFR capabilities wherein resources with static droop characteristics provide PFR but then use a direct load frequency control approach for secondary control. These combined approaches have the advantage of fast control followed by achieving a dispatched set point on a slower timescale, but they still have significant uncertainty and nonoptimal dispatch because their governing PFR control parameters are still updated on a slower timescale that does not compensate for changes in load and DER occurring on a faster timescale. In addition, these approaches are not fast enough for FFR. In our previous work [31], we experimentally demonstrated optimal coordination of four loads and an inverter within a single building for FFR within 142 ms.

C. Contribution and Organization

This two-part paper makes the following contributions:

- 1) We develop a first-of-its-kind control solution that optimally coordinates tens of thousands of deferrable loads and DERs across a distribution network to provide bulk power system-scale FFR that is dispatchable and accurate. Compared to existing frequency response solutions, the developed control platform is able to optimally coordinate a much larger number of net-load resources (up to 40,000 is shown), incorporate fast-changing BTM resources, and accurately meet a specific aggregate power set point while also achieving dispatch on a faster timescale (<500 ms) suitable for FFR.
- 2) We provide an analytical proof of the atomic control algorithm's optimality and then verify near-optimal (within 0.1%) performance of the algorithm as it is scaled to configurations with up to 40,000 net-load resources via Monte Carlo-style simulation. In all cases, dispatch solutions are achieved faster than traditional mixed-integer linear program (MILP) optimization approaches and sufficiently fast to support coordinated FFR within 500 ms.

- 3) We present a solution for determining the optimal amount of headroom to operate solar inverters with to minimize opportunity cost while ensuring the FFR response viability of a building with the inverter and deferrable loads. A practical optimal headroom algorithm is given and analytically proven to match a bi-level optimization-based approach.

- 4) The entire platform is implemented and experimentally verified using a unique power hardware-in-the-loop (PHIL) demonstration including more than 100 powered loads and DERs connected to a real-world distribution network model from Colorado, USA. Test cases demonstrate the approach at multiple scales, including a full-scale demonstration with more than 10,000 deferrable loads and DERs.

This two-part paper is organized as follows: Part I presents an overview of the control platform in Section II, a definition of the net-load unit (NLU) building block in Section III, an algorithm for determining optimal NLU inverter headroom that is developed with analytical guarantees in Section IV, algorithms for optimal FFR dispatch from a single NLU in Section V, and algorithms for optimal FFR dispatch from multiple NLU and net-load groups (NLGs) that are developed and validated via simulation results in Section VI. In Part II [32], the entire platform is implemented and experimentally verified with multiple test cases.

II. PLATFORM FOR FAST FREQUENCY RESPONSE FROM OPTIMALLY-COORDINATED NET-LOAD RESOURCES

A. Objective and Requirements

The objective of this work is to achieve a platform that can optimally coordinate tens of thousands of net-load resources connected in an electric distribution network as a VPP to provide FFR ancillary services to the bulk power system. In particular, the platform should:

- meet general FFR service performance requirements (i.e., dispatchable, ≥ 1 -MW capacity, response within 500 ms [includes all communications delays], and achieve set point within 1%)
- coordinate the operating power level and on/off status of tens of thousands of net-load resources to achieve the lowest global cost, while meeting the prescribed FFR set point for the VPP and respecting all resource limits
- incorporate fast-changing BTM resources, such as deferrable loads and solar inverters, in its optimal dispatch
- provide online control for participating solar inverters to operate each inverter with an optimal amount of headroom to ensure the FFR response viability of a building with the inverter and deferrable loads, while minimizing opportunity cost
- allow for real-time user input of resource price and availability; individual devices can participate at will
- preserve user privacy and allow for a variety of communication protocols

B. Solution Overview

As shown in Fig. 1(a), the developed platform leverages a hierarchical structure that is scalable, preserves user privacy, enhances reliability, and provides optimal coordination at selectable levels. At the lowest level, individual DERs and flexible loads are organized into NLUs wherein a net-load unit

controller (NLUC) manages participating net-load resources. Loads are controlled through plug- or circuit-level load sensing and control interfaces (LSCIs), which provide high-speed (half ac cycle) power measurement and on/off actuation capabilities (see [31] for more details). Each NLUC also implements an optimal NLU inverter headroom algorithm (see Section IV) to operate constituent solar inverters to minimize opportunity cost while ensuring the FFR response viability of the NLU. NLUs, typically representing individual buildings, are organized into NLGs, which generally represent all the NLUs connected to a single distribution transformer and are managed by a net-load group controller (NLGC).

For global coordination, a distribution system master controller (DSMC) can optimally dispatch all NLUs across a distribution network according to FFR set points and frequency trigger levels provided by the balancing authority or higher-level aggregator. In this configuration, the computational burden of optimal dispatch calculation is distributed across the NLGCs, allowing nearly constant dispatch calculation times regardless of how many NLGs are participating. In this hierarchical structure, user privacy is preserved as controllers can communicate only one layer up or down and only minimal device information, such as operating power and capacity limits, is shared (e.g., NLGs cannot directly address devices and do not have knowledge of specific devices). An independent communications network is used at each layer, enabling specific communications protocols or security measures to be implemented at each layer as needed. A communications relay is used to interface the NLGCs and the DSMC—this provides a flexible, high-performance method to communicate information between each NLGC and the DSMC in a redundant, plug-and-play fashion without the DSMC needing to individually address each NLGC. Building owners interact with the NLUC by setting unit costs and enabling/disabling dispatch for each resource.

C. Method of Operation

Optimal net-load resource dispatch can be achieved within a single NLU, an NLG with multiple buildings, or across the entire distribution network when coordinating NLGCs using DSMC. Algorithms and performance evaluations for optimal dispatch at each of these three scales are given in the following sections. In this section, we describe the method of operation for the most interesting case—optimal FFR dispatch of a distribution feeder-scale VPP with tens of thousands of net-load resources. FFR dispatch within a single NLU or individual NLG are subcases that are described in more detail in Section V and Section VI, respectively. Explain what happens if there is a fault in the communication system.

1) *Pre-event Actions:* Before a frequency event occurs, continuous action within the VPP prepares it for effective FFR response. Within each NLU, each device controller continuously updates the NLUC with the latest device measurements and operating limits. Each NLUC uses local device measurements as inputs to the optimal inverter headroom algorithm (see Section IV) and continuously updates the operating point of its constituent solar inverter. Summary power measurement, capacity, and cost information is aggregated at each NLUC and NLGC before being sent to the DSMC. The DSMC uses

this summary information to inform the FFR bids that it places with the balancing authority or other higher-level aggregator. In response, the balancing authority sends the DSMC the FFR set point it is committed to provide and the associated frequency trigger levels outside of which this response should be activated.

2) *Event Dispatch Actions:* The sequence of actions to execute a FFR of the full VPP from the instant of a grid frequency contingency event is shown in Fig. 1(b). Most commonly, frequency is sensed locally within the distribution network (e.g., via relay or a DER with high-speed frequency reporting capabilities), and used to, when appropriate, send an event trigger to the DSMC to begin the dispatch sequence. The DSMC can also act upon a direct event trigger from the balancing authority. In the first step of the dispatch sequence, the DSMC sends a request to all participating NLGCs (determined by those that have sent an updated status measurement recently) to provide the set of best net-load response options within their respective NLGs. The NLGCs then calculate all net-load response options within their NLG when considering all NLUs (and their constituent devices) that have sent updated status information recently using Mthd. 1 (see Mthd. descriptions in Section VI-B and Fig. 3); ii) downselect to the best (highest likelihood of being included in the global optimal dispatch solution) net-load response options using Mthd. 4; and iii) send these best net-load response options back to the DSMC. The DSMC then gathers the best net-load response options from all participating NLGCs, runs an algorithm (Mthd. 5) to downselect to the net-load response options that result in the optimal (lowest cost while meeting the given VPP power set point) total dispatch action, and sends out the final net-load dispatch actions to each NLGC. Finally, commands are routed through each NLGC and NLUC down to the individual device controllers which execute them. The megawatt-scale FFR of the distribution network-scale VPP can then be measured (and likely verified a meter by the balancing authority) at the distribution feeder head.

With this sequence of operations, the computational burden of determining the global optimal dispatch actions is shared by all NLGCs, which all calculate their response options in parallel, and the DSMC. This allows for excellent scalability, maintaining fast computation regardless of how many NLGs are participating. In centralized solutions, the complexity of the dispatch problem grows exponentially with the number of participating resources; our solution avoids this curse of dimensionality by distributing computation across the NLGCs and then running a computationally efficient algorithm for final coordination at the DSMC. This distributed solution does require communications between each layer in the hierarchical control solution. Thus, communications delay times, in addition to other practical factors such as the actuation times of the end devices, must be accounted for in achieving an overall FFR dispatch within 500 ms. As shown in Section VI-C2 and in the experimental results of Part II [32], we develop the hierarchical dispatch algorithm to reliably compute a solution within 150-200 ms (see Fig. 6(b)), leaving over 300 ms for communications and device actuation. Communications reliability is important, though the hierarchical control solution's

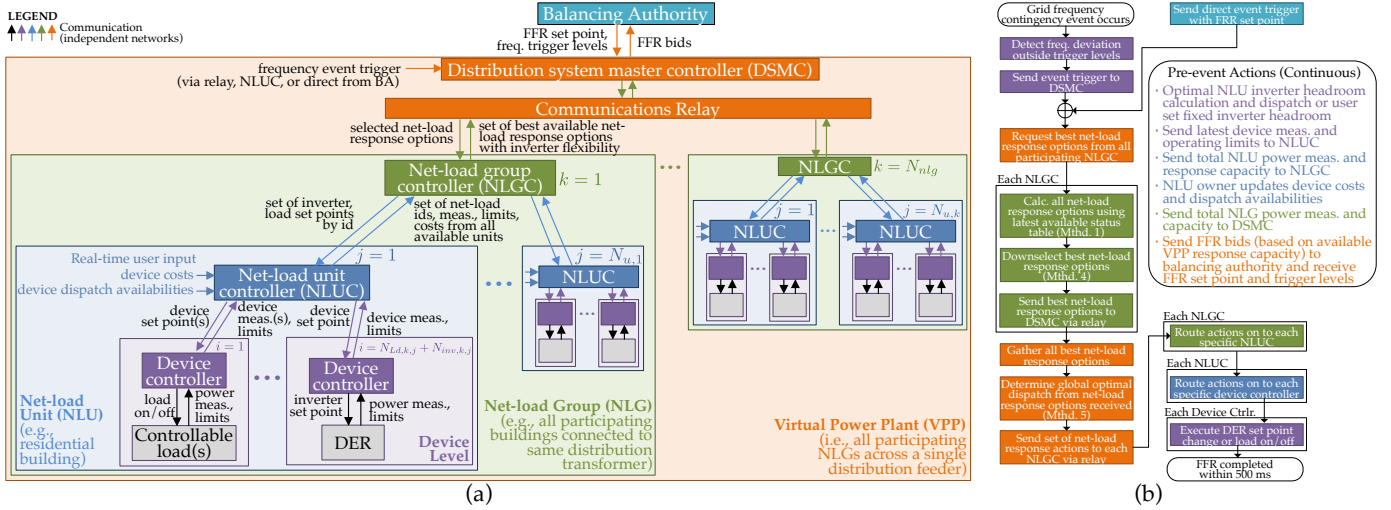


Fig. 1. Hierarchical optimal net-load coordination for FFR approach. (a) Block diagram showing connections and key data passed between each of the four layers. (b) Flow chart indicating the sequence of steps to coordinate all net-loads into a VPP and provide MW-scale FFR in response to an event trigger.

plug and play design with fast (ms timescale) status updates greatly mitigates the potential impacts of any communications faults; participating resources are only included in the final dispatch calculation if those resources have very recently (ms timescale) provided a status update indicating their online nature and willingness to participate. In the unlikely event that a particular resource's communications drop offline in the 150-200 ms between their last status communication and the conclusion of the dispatch calculation, the total FFR power set point will not be met exactly; however, we aim to meet the the FFR power set point within an acceptable 1% tolerance.

III. NET-LOAD UNIT DEFINITION

At its most basic level, an NLU (e.g., see the bottom of Fig. 2) contains nondeferrable load with aggregate active power consumption, P_L ; N_{Ld} deferrable loads with active power consumptions, $\{P_{Ld,i}\}_{i=1}^{N_{Ld}}$; an inverter that outputs active power, P_{inv} ; and active power consumption, P_G (i.e., $P_G > 0$ represents power consumed by the NLU), from the NLU's grid connection. Note that power consumed or supplied at each point throughout this problem formulation are all time-varying quantities; however, in most cases, the time dependence (t) is dropped for ease of notation. All $P_{Ld,i}$ are deferrable such that their final power consumption is $\hat{P}_{Ld,i} = \bar{x}_i P_{Ld,i} = (1 - x_i) P_{Ld,i}$ after being issued an on/off command, $x_i \in \{0, 1\}$. That is, $x_i = 1$ denotes that the load has been deferred or switched off. The inverter is initially operating at power $P_{inv}(t_0) = P_{inv,base}$ and has operating limits $0 \leq P_{inv} \leq P_{inv}^{max}$. Upon command, the inverter can change its output power to $P_{inv} = P_{inv,base} + P_{inv\Delta}$, where $P_{inv\Delta}$ is the inverter's change in power from $P_{inv}(t_0)$. The inverter's change in power is limited: $P_{inv\Delta,dn} \leq P_{inv\Delta} \leq P_{inv\Delta,up}$, where $-P_{inv,base} \leq P_{inv\Delta,dn} \leq 0$ limits how much the inverter can decrease its power output, and $0 \leq P_{inv\Delta,up} \leq (P_{inv}^{max} - P_{inv,base})$ limits how much the inverter can increase its power output. Throughout this work, we assume $P_{inv\Delta,dn} = -P_{inv,base}$ for simplicity because curtailment of inverter output is generally not a significant constraint. $P_{inv\Delta,up}$ is known as the inverter's headroom and will be determined in real time using the algorithm developed

in Section IV. By Kirchoff's laws, the power balance on the net-load system is given by: $\sum_{i=1}^{N_{Ld}} \hat{P}_{Ld,i} + P_L - P_{inv} = P_G$.

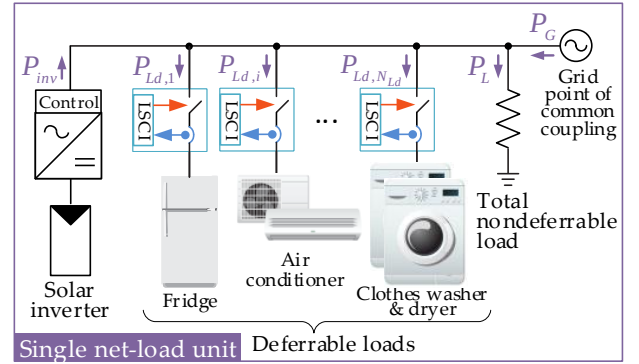


Fig. 2. One-line diagram of an example net-load unit. DERs, such as solar inverters, and controllable loads can be controlled directly with real-valued power set points. Loads can be configured as fast-acting deferrable loads by interconnecting them through a load sensing and control interface (LSCI). Often, a portion of the NLU's load cannot be controlled or deferred.

The NLU owner determines in advance of the frequency contingency event which loads are deferrable vs. nondeferrable and also assigns a participation unit cost, c_i , for each deferrable load as well as the inverter (c_{inv}) such that the owner will be compensated a total of $c_{tot} = \sum_{i=1}^{N_{Ld}} x_i(t) \cdot c_i \cdot P_{Ld,i}(t_0) + c_{inv} |P_{inv}(t) - P_{inv}(t_0)|$ after providing FFR. These determinations are updated by the building manager in real time so that any frequency event response considers the owner's latest preferences. This simplistic compensation scheme is meant to replicate key aspects of market participation, including users determining if they want to participate and at what price, determining total cost for an aggregated response action, and compensating users. Each NLU periodically (and again immediately after receiving an event response trigger) sends its measured P_G , $\{\hat{P}_{Ld,i}\}_{i=1}^{N_{Ld}}$, P_{inv} , $P_{inv\Delta,dn}$, and $P_{inv\Delta,up}$ to the NLGC. The NLGC uses this information to determine an optimal $\Delta P_G^* = P_G(t) - P_G(t_0)$ for each NLU shortly after an FFR event is triggered and subsequently verify the response performance of each NLU. The NLGC will ensure that NLU's are provided with ΔP_G^* set points within their operating capability; that is, $-(P_{inv\Delta,up} + \sum_{i=1}^{N_{Ld}} \hat{P}_{Ld,i}) \leq \Delta P_G^* \leq P_{inv\Delta,dn}$.

IV. OPTIMAL NLU INVERTER HEADROOM

Here, we consider the problem of determining the optimal amount of headroom, $P_{inv\Delta,up}$, that a solar inverter within an NLU should maintain prior to an FFR underfrequency response dispatch event. This issue is important in the context of the overall NLU optimal dispatch problem (see Section V-A) because the inverter headroom is set in advance of a dispatch event and determines what values of dispatch command ΔP_G^* can be achieved by the NLU. Insufficient inverter headroom can make the single NLU optimal dispatch problem infeasible or result in a sub-optimal dispatch solution.

To illustrate the need for headroom, consider the simple case where an NLU has $N_{Ld} = 2$ deferrable loads, $P_{Ld}(t_0) = [100, 300]W$, which are both initially on ($x(t_0) = [0, 0]$), and one solar inverter. A grid underfrequency event occurs and shortly later the NLU is given a dispatch command, $\Delta P_G^* = P_G(t) - P_G(t_0)$, to consume less grid power ($\Delta P_G^* < 0$). To meet this dispatch command, the NLU must curtail load(s) and/or modify inverter output. Controlling the on/off state of the N_{Ld} deferrable loads in the NLU achieves $2^{N_{Ld}}$ possible NLU power responses ($P_{LdTp} = \{0, 100, 300, 400\}W$ for this case). Unless $|\Delta P_G^*| \in P_{LdTp}$, a change in inverter output will be required for the NLU's total grid power response, $\Delta P_G = -\sum_{i=1}^{N_{Ld}} x_i P_{Ld,i} - P_{inv\Delta}$, to meet the dispatch command, ΔP_G^* , exactly, which is a condition of NLU dispatch problem feasibility (see Section V-A). For example, if a command of $\Delta P_G^* = -125W$ was given, there are two possible solutions: (i) loads are curtailed to $\sum_{i=1}^{N_{Ld}} x_i P_{Ld,i} = 100W$, requiring $P_{inv\Delta} = 25W$; or (ii) loads are curtailed to $\sum_{i=1}^{N_{Ld}} x_i P_{Ld,i} = 300W$, requiring $P_{inv\Delta} = -175W$. Solution (i) requires both less curtailment of load and less total change in inverter power, resulting in a lower cost dispatch solution, but this requires the solar inverter to increase its power output. However, this solution would not be feasible if the solar inverter is already operating at its maximum power ($P_{inv} = P_{inv}^{max}$, which is a typical goal of solar inverters) with no headroom available. Solution (ii) is a higher cost solution, but it does not require any upward headroom. For this example ($\Delta P_G^* = -125W$), having ensured inverter headroom of $P_{inv\Delta,up} \geq 25W$ in advance of the dispatch event would have allowed solution (i) to be feasible and the NLU dispatch problem to achieve the global optimal result. If insufficient inverter headroom ($P_{inv\Delta,up} < 25W$) is available in advance, then the NLU dispatch problem would be forced to use solution (ii), a less optimal result. This illustrates the needed headroom for a particular ΔP_G^* . The value of ΔP_G^* is not known in advance, however, and therefore to ensure feasibility for any possible $-(P_{inv\Delta,up} + \sum_{i=1}^{N_{Ld}} \hat{P}_{Ld,i}) \leq \Delta P_G^* \leq P_{inv\Delta,dn}$ given to the NLU, the optimal inverter headroom problem must consider all possible ΔP_G^* . In the simple case illustrated, the global maximum headroom required occurs for $\Delta P_G^* = -200W$, where $P_{inv\Delta,up} = 100W$ is required.

A. Problem Formulation

The optimal NLU inverter headroom problem objective is to ensure sufficient headroom, $P_{inv\Delta,up}$, such that the optimal result of the subsequent NLU optimal dispatch problem at any $-(P_{inv\Delta,up} + \sum_{i=1}^{N_{Ld}} \hat{P}_{Ld,i}) \leq \Delta P_G^* \leq 0$ is feasible

while simultaneously minimizing $P_{inv\Delta,up}$ to minimize the opportunity cost of operating the solar inverter off of its maximum power point. This can be accomplished by solving the following bi-level optimization problem:

$$\alpha = \max_{\Delta P_G^*} f_2 \quad (1)$$

$$\text{s.t.} \quad -(P_{inv}^{max} + \sum_{i=1}^{N_{Ld}} P_{Ld,i}) \leq \Delta P_G^* \leq 0 \quad (2)$$

$$\text{where } f_2 = \min_{\tilde{x}_i} \Delta P_G^* + \sum_{i=1}^{N_{Ld}} \tilde{x}_i P_{Ld,i} \quad (3)$$

$$\text{s.t.} \quad \tilde{x}_i \in \{0, 1\} \forall i \quad (4)$$

and then setting $P_{inv\Delta,up} = \min(\alpha, P_{inv}^{max})$. Note that if $\alpha > P_{inv}^{max}$, then the NLU will be unable to meet every possible given ΔP_G^* , though setting $P_{inv\Delta,up} = P_{inv}^{max}$ will give the NLU the greatest probability of meeting all possible ΔP_G^* . In this bi-level optimization problem, the inner problem (3)–(4) determines the inverter headroom required at the NLU dispatch solution for a given ΔP_G^* when utilizing the closest possible load combination, and the outer problem (1)–(2) searches for the maximum inverter headroom result across all ΔP_G^* possible for the NLU's capacities. Here, (2) considers an upper limit of $\Delta P_G^* \leq 0$ because we focus on underfrequency events where inverter upward headroom is utilized.

B. Solution

Here, we develop a fast algorithm that solves the optimal inverter headroom problem (1)–(4) and without requiring a bi-level optimization solver. A simple recurrence-based implementation is given in Algorithm 1. A proof that this algorithm achieves the optimal result is given in Theorem 1.

Theorem 1. Let $P_{Lds} = \langle P_{Ld}, \leq \rangle$, where P_{Lds} is an ordered version of P_{Ld} with order given by $\leq$. Then, the result, α , of the optimal inverter headroom problem (1)–(4) can be obtained as $\alpha = \max_{1 \leq i \leq N_{Ld}} (P_{Lds,i} - \sum_{n=1}^{i-1} P_{Lds,n})/2$.

Proof: See Appendix A. ■

Algorithm 1 Optimal NLU inverter headroom recurrence

```

1: Require  $P_{Lds}, P_{inv}^{max}$ 
2: function maxLD( $lds, ldSum, n$ )
3:   if  $n == (\text{length}(lds) - 1)$  then
4:     return  $lds[n] - ldSum$ 
5:   else
6:     return  $\max((lds[n] - ldSum), \text{maxLD}(lds, ldSum + lds[n], n + 1))$ 
7:   end if
8: end function
9:
10: Begin
11:  $P_{inv\Delta,up} = \min(\text{maxLD}(\text{sorted}(P_{Lds}), 0, 0)/2, P_{inv}^{max})$ 
12: return  $P_{inv\Delta,up}$ 

```

V. OPTIMAL FAST FREQUENCY RESPONSE FROM A SINGLE NET-LOAD UNIT

A. Dispatch Problem Formulation

Prior to an FFR dispatch event occurring, a single NLU has initial conditions of $x_i = 0$ for all i (all participating deferrable loads are on) and $P_{inv} = P_{inv,base}$ with headroom $P_{inv\Delta,up}$. For simplicity in this single-unit problem, resources are all considered to have unit cost (more complex cases with varying cost are considered in Section VI). The NLU is given a commanded change in grid active power consumption,

ΔP_G , to follow to provide the requested FFR. The objective of the NLU is to track an active power consumption at its grid connection of $P_G^* = P_G(t_0) + \Delta P_G$, while minimizing the amount of deferrable load that must be curtailed and the change in inverter power required. This objective is given by:

$$\mu = \min_{x, P_{inv\Delta}} \sum_{i=1}^{N_{Ld}} x_i P_{Ld,i} \quad (5)$$

$$\text{subject to } \sum_{i=1}^{N_{Ld}} (1 - x_i) P_{Ld,i} + P_L - P_{inv,base} \quad (6)$$

$$- P_{inv\Delta} = P_G^* \quad (7)$$

$$P_{inv\Delta,dn} \leq P_{inv\Delta} \leq P_{inv\Delta,up} \quad (8)$$

$$x_i \in \{0, 1\} \forall i = 1 \dots N_{Ld} \quad (8)$$

Solving this optimization problem yields the solution $(x^\mu, P_{inv\Delta}^\mu)$. Let the feasible set for this optimization problem be $\Omega^\mu := \{(x, P_{inv\Delta}) \mid (6)-(8) \text{ are met}\}$.

B. Solution

The dispatch problem (5)–(8), with its mix of integer (e.g., x_i) and real-valued (e.g., $P_{inv\Delta}$) decision variables, can be solved using an MILP. This approach is possibly tractable for the single NLU dispatch problem with a modest number of decision variables; however, it is too slow to support dispatch on FFR timescales for larger problems (e.g., dispatching 1000s of net-load resources), as shown in Section VI-C2. Instead, we develop a simpler approach that obtains the same optimization result for a single NLU (and nearly optimal for the aggregated case) but with faster performance and better scalability.

The dispatch problem (5)–(8) can be simplified to a binary programming problem with relaxed constraint (6), followed by a single variable update, as shown in (13). The advantage is that it can be solved using a dynamic program, and we show that it can be solved faster using this method than with a standard MILP. The simplified approach is formulated as:

$$\nu = \min_x \sum_{i=1}^{N_{Ld}} x_i P_{Ld,i} \quad (9)$$

$$\text{subject to } \sum_{i=1}^{N_{Ld}} (1 - x_i) P_{Ld,i} + P_L - P_{inv,base} \quad (10)$$

$$- P_{inv\Delta,up} \leq P_G^* \quad (11)$$

$$- P_G^* \geq P_{inv\Delta,dn} \quad (12)$$

$$x_i \in \{0, 1\} \forall i = 1 \dots N_{Ld} \quad (12)$$

Let x^ν be the solution to this optimization problem. Finally, to complete the simplified net-load dispatch, $P_{inv\Delta}$ is set so that P_G^* is met exactly:

$$P_{inv\Delta}^\nu := \sum_{i=1}^{N_{Ld}} (1 - x_i^\nu) P_{Ld,i} - P_{inv,base} - P_G^*. \quad (13)$$

Let the feasible set for this optimization problem be $\Omega^\nu := \{(x, P_{inv\Delta}) \mid (10)-(12) \text{ are met}\}$. We prove that this simplified approach yields the same optimal solution as the optimal MILP approach in Section V-C.

C. Solution Equivalency Proof

Theorem 2. $\nu = \mu$. That is, the simplified net-load dispatch problem (9)–(13) yields the same solution as the net-load dispatch optimization problem (5)–(8).

Proof: See Appendix B. ■

VI. OPTIMAL FAST FREQUENCY RESPONSE FROM MULTI-BUILDING NET-LOAD GROUPS

A. Problem Formulation

Optimal coordination of net-load resources for FFR can occur for multiple buildings within a single NLGC or, for globally optimal results, via coordination of all NLGCs across the distribution system managed by a single DSMC. Here, we consider the more general problem formulation: a global, distribution feeder-wide dispatch of N_{nlg} NLGCs for FFR. The k^{th} NLGC coordinates $N_{u,k}$ NLUs and the j^{th} NLU of the k^{th} NLGC has $N_{Ld,k,j}$ deferrable loads, which are all active at t_0 . The i^{th} deferrable load and the inverter of the j^{th} NLU under the k^{th} NLGC have cost $c_{kji} > 0$ and $c_{inv,kj} > 0$, respectively. The i^{th} deferrable load consumes power $P_{Ld,kji}$, and the inverter produces power $P_{inv,k,j} = P_{inv,k,j,base} + P_{inv\Delta,k,j}$. The bounds on the change in the inverter power allowable are constrained by the headroom given by $P_{inv\Delta,up,k,j}$ as an upper bound and $P_{inv\Delta,dn,k,j}$ as a lower bound. The group of NLGCs, managed by a single DSMC, is given a commanded change in total active power, ΔP_{Gg} , that the aggregated response of the NLUs should collectively follow to provide the requested FFR. The objective of the coordinated NLGs is to track a total grid active power consumption $P_{Gg}^* = P_{Gg}(t_{dg}) + \Delta P_{Gg}$ (where $P_{Gg} = \sum_{k=1}^{N_{nlg}} \sum_{j=1}^{N_{u,k}} P_{G,k,j}$ and $P_{G,k,j}$ is the active power consumed at the grid connection of the j^{th} NLU in the k^{th} NLG) while minimizing the total cost. This objective can be formalized as:

$$\min_{x_{k,j}, P_{inv\Delta,k,j}} \sum_{k=1}^{N_{nlg}} \left(\sum_{j=1}^{N_{u,k}} \left(\sum_{i=1}^{N_{Ld,k,j}} x_{kji} c_{kji} P_{Ld,kji} + c_{inv,k,j} |P_{inv,k,j}(t) - P_{inv,k,j}(t_0)| \right) \right) \quad (14)$$

$$\text{subject to } \sum_{k=1}^{N_{nlg}} \left(\sum_{j=1}^{N_{u,k}} \left(\sum_{i=1}^{N_{Ld,k,j}} (1 - x_{kji}) P_{Ld,kji} + P_{L,k,j} - P_{inv,k,j,base} - P_{inv\Delta,k,j} \right) \right) = P_{Gg}^* \quad (15)$$

$$P_{inv\Delta,dn,k,j} \leq P_{inv\Delta,k,j} \leq P_{inv\Delta,up,k,j} \quad (16)$$

$$\forall j = 1 \dots N_{u,k}, k = 1 \dots N_{nlg} \quad (17)$$

$$P_{inv\Delta,k,j} = P_{inv,k,j}(t) - P_{inv,k,j}(t_0) \quad (17)$$

$$\forall j = 1 \dots N_{u,k}, k = 1 \dots N_{nlg} \quad (18)$$

$$x_{kji} \in \{0, 1\} \quad (18)$$

$$\forall i = 1 \dots N_{Ld,k,j}, j = 1 \dots N_{u,k}, k = 1 \dots N_{nlg} \quad (18)$$

Here, $P_{inv\Delta,k,j}$ is the determined change in inverter power for the j^{th} NLU; $P_{L,k,j}$ is the aggregate nondeferrable load active power for the j^{th} NLU; and $P_{Ld,kji}$ is the active power of the i^{th} deferrable load in the j^{th} NLU. Solving this optimization problem yields the solution $(x^\rho, P_{inv\Delta}^\rho)$.

B. Solution

Similar to the single-NLU dispatch problem, the optimal dispatch problem (14)–(18) can be solved directly using an MILP; however, even when using state-of-the-art optimization solvers, solutions are not always possible for large problems because of limits on the number of variables or max recursion depth that can be achieved. When solutions are possible, solution times are often too slow to support dispatch on FFR timescales, as shown in Section VI-C2. Thus, we develop a hierarchical approach that builds on the simplified NLU dispatch approach of Section V-B and distributes computation across NLG controllers to calculate dispatch results on a timescale fast enough for FFR while maintaining near-optimal dispatch results. This control approach can be applied within a single NLG or for all NLGs across a distribution feeder.

Methods for both scales are summarized in Fig. 3. Here, A_{sel} is a matrix of length $|A_{sel}| = 2^{N_o}$ ($N_o = 20$ was used herein), where each row is the N_o -bit binary representation of that row index; $interp1d(y, f(x), x)$ represents a 1-D linear interpolation of function $f(x)$ at point y ; $\dagger$ indicates a wild card for the two inputs to that block (i.e., the function with $\dagger$ should be applied to both inputs independently); $*$ indicates a set point or optimal point; and ΔP^* is the desired change in global net-load power consumption for the FFR event. Optimal dispatch within a single NLGC does not involve the DSMC. At the conclusion of a global dispatch, which involves NLGCs and the DSMC, each NLGC with id, z , is sent its $N_{r,z}$ corresponding dispatch rows of the form $\{\Delta P_{L,m}, \Delta P_{inv,m}, \Delta P_{Tnl,m}, i_m, z\}_{m=1}^{N_{r,z}}$. Each NLGC then compares these dispatch rows with the array of net-load options it stored at the event's initiation and executes the appropriate load curtailments and inverter set point changes.

Optimal FFR dispatch of a single NLG is accomplished using Mthd. 1-3. In Mthd. 1, the net-load response capacities (change in load if deferred or total possible change in inverter power) and their associated unit participation costs from all devices within all NLUs in the NLG are used to calculate $|A_{sel}| = 2^{N_o}$ possible net-load response options from the N_o lowest-cost net-loads. For single NLG dispatch, global optimality of the final dispatch solution can be guaranteed if $N_{Tld} = \sum_j N_{Ld,j} \leq N_o$. However, as demonstrated in Section VI-C2 and the experimental results of Part II [32], near optimal (within $<0.2\%$) dispatch performance can still be achieved for $N_o < N_{Tld} \leq 2N_o$. The next step in single NLG dispatch is to employ Mthd. 2 to get the total inverter flexibility profile. In short, Mthd. 2 takes the set of downward headroom, upward headroom, and unit cost for all inverters in the NLG to build a single array representing every possible total inverter response (upward and downward) and the associated total cost for each response. As shown in Fig. 4(a), this essentially adds up the linear capacity vs. cost characteristics of each inverter to represent a single large inverter with a piecewise-linear inverter capacity vs. total cost characteristic. Finally, Mthd. 3 determines the optimal combination of net-load response (from Mthd. 1) and inverter flexibility (from Mthd. 2) to determine the optimal dispatch decisions to send to individual net-loads within the NLG. As shown in Fig. 4(b), this method

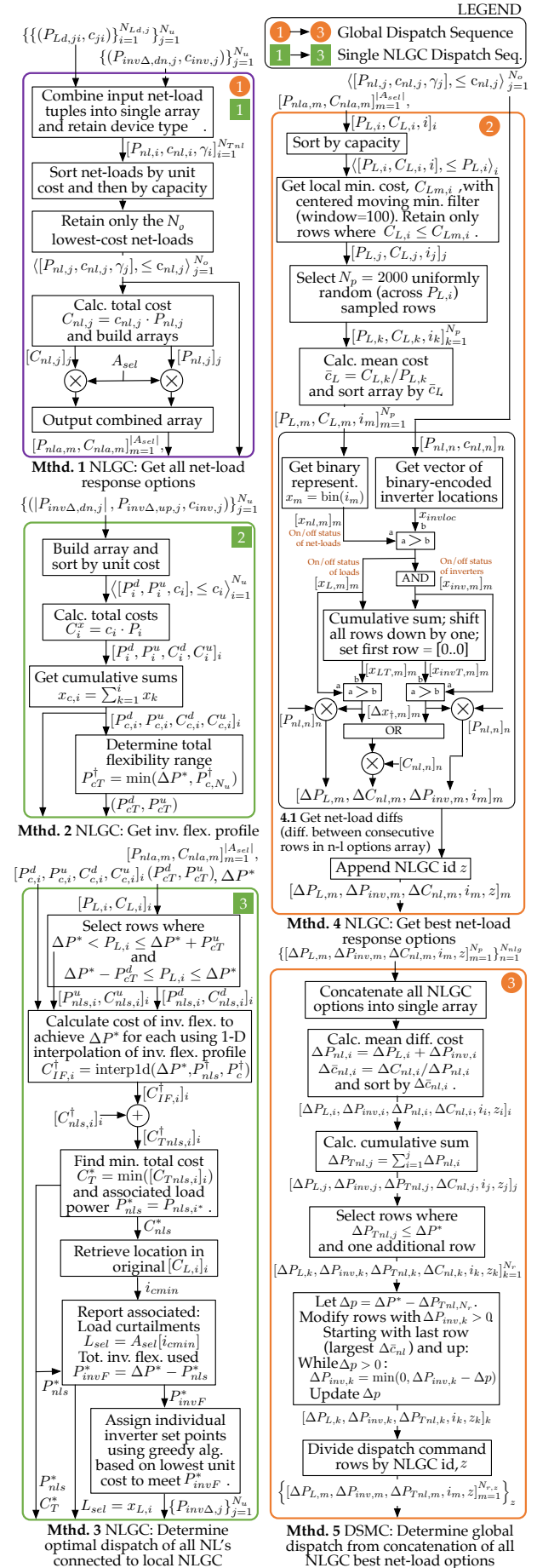


Fig. 3. Hierarchical optimal net-load dispatch algorithms

considers all net-load response options within the total inverter flexibility range of the FFR setpoint (ΔP_G^* in Fig. 4(b)) and then determines which discrete net-load response combined with inverter flexibility profile achieves the FFR setpoint at the lowest total cost. This follows the same methodology as the proven solution in Section V—first select individual net-load resource responses (here, both load and full inverter capacities are considered) using binary variables to obtain solutions close to the FFR set point, and then use the inverter (here, the total inverter flexibility profile made of up many individual inverters) to obtain the optimal solution.

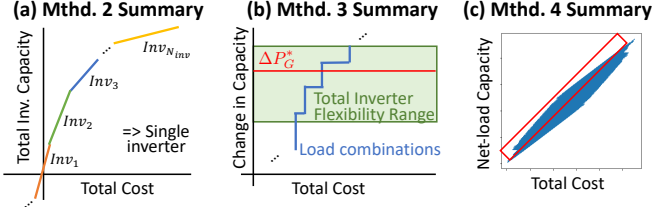


Fig. 4. Summary and intuition of Mthds. 2-4 from Fig. 3: (a) the linear cost vs. capacity characteristics of many smaller inverters can be summed up to represent a single large inverter with piecewise-linear characteristics; (b) the global optimal dispatch solution is obtained by considering all discrete net-load response options around the FFR capacity and then using total inverter flexibility; and (c) the “best” net-load response options are the lowest-cost options across the full range of net-load response capacities.

Optimal FFR dispatch of the full VPP, including multiple NLGs and their constituent NLUs and devices, is accomplished using the sequence of steps shown in Fig. 1(b). As part of this sequence, each participating NLGC calculates all its net-load response options using Mthd. 1 and then, as illustrated in Fig. 1(c), downselects this set of options to the “best” options, representing the lowest-cost options across all net-load response capacities possible from the NLG, using a moving minimum filter. The exact quantity of best options can be set by the user—in practice, we found $N_p = 2000$ to work well. These best options are then sorted by lowest mean cost and then differences (“net-load diffs” in Fig. 3; includes difference in net power and inverter power) between consecutive options are calculated so that only unique net-load actions (e.g., one net-load response option could be a superset of another, in that all net-load actions are the same, except for one additional net-load action is present in the first) are considered and this final set sent back to the DSMC. The mapping of absolute net-load actions to the differences provided to the DSMC are stored at each NLGC so that the appropriate load curtailments and inverter set point changes can be determined after receiving a list of final selected net-load options (based on differences) from the DSMC. To determine the final optimal dispatch solution, the DSMC, uses Mthd. 5 to combine the best net-load options from all participating NLGs (output of Mthd. 4) together, sort by lowest mean cost, and then uses a greedy algorithm to select the lowest-cost discrete net-load response options that are part of the optimal dispatch solution. Finally, the lowest-cost inverter flexibility from the not-selected net-load options is assigned until the FFR set point is achieved.

C. Performance Verification

The hierarchical approach developed employs a heuristic method that can be executed significantly faster than optimization-based approaches, trading complete optimality for speed. Thus, it isn’t tractable to analytically prove that the hierarchical approach will achieve the optimal dispatch result 100% of the time. Instead, we evaluate performance using Monte Carlo-style simulation and determine distributions of probability vs. percent optimality (measured via solution cost) and probability vs. solution time. To do so, all simulation runs solve a multi-NLU dispatch problem twice—once using the hierarchical approach implemented in Python and again using an MILP solution deployed on a state-of-the-art, commercial (Gurobi) mixed integer program solver. The MILP solution provides the reference global optimal solution cost and the solution times of both approaches are compared. Both simulation types use the same scenario inputs and are executed on the same computer with a six-core 2.9 GHz Intel Core i9 processor. The hierarchical approach is designed to use only one processor core (to allow for greater portability to a variety of embedded platforms) and is implemented as such, whereas the MILP solution is allowed to use all six processor cores.

1) *Single Net-load Group Case:* Simulations were performed for optimal FFR dispatch of a single NLG with $N_u = 10$ NLUs; each had 4 loads and 1 inverter, for a total of 50 time-varying net-load resources to be dispatched. More than 2 million simulation runs with deferrable load capacities uniformly distributed in the interval $P_{Ld,j,i} \in [100, 6000] W$; inverters with uniformly distributed capacity $P_{inv,j}^{max} \in [2000, 5000] W$ and headroom $P_{inv,j,up} \in [0, 2000] W$; uniformly distributed NLU costs $c_{ji} \in [2, 10]$; and an assigned NLG participation factor (the percentage of the NLG’s total net grid power consumption that is dispatched) $\eta \in (0, 45]\%$ were run. Solution optimality results in Table I confirm that, on average, the hierarchical approach achieves an optimal dispatch decision with a solution cost very close (within 0.00034%) to the optimal, that solution cost is optimal 99.997% of the time, that the hierarchical approach has a mean solution time of 52.2 ms, and that there is a 99.992% probability of achieving a solution within 120 ms.

TABLE I.
SINGLE NLG CASE: SUMMARY SIMULATION RESULTS

Number of simulation runs	2,007,890
Mean % difference in solution cost, HA vs. MILP	0.00034%
Probability of HA solution cost within 1% of optimal	99.997%
Probability of HA solution cost within 0.1% of optimal	99.912%
Probability of HA solution cost within 0.01% of optimal	99.842%
Mean solution time, HA	52.2 ms
Mean solution time, MILP	107.1 ms
Probability of HA solution time < 120 ms	99.992%
Probability of HA solution time < 100 ms	99.962%
Probability of HA solution time < 80 ms	98.935%

2) *Full Feeder Case:* The optimal FFR dispatch problem for a full-scale electric distribution system feeder with a varying number of participating NLGs was simulated in 9,920 simulation runs to evaluate the performance of the proposed hierarchical approach versus the MILP solution and determine how their performances scaled as the number of participating net-load resources and the total feeder participation factor

varied. In the simulation, each NLG included $N_u = 10$ NLUs, each with 4 loads and 1 inverter participating, connected to a single distribution transformer and coordinated by an NLGC. The simulations included 155 runs at each of 64 possible combinations of the number of NLGs participating, $N_{NLG} = [100, 200, 300, 400, 500, 600, 700, 800]$, and the total feeder dispatch participation factor, $\eta = [10, 15, 20, 25, 30, 35, 40, 45]\%$. Each simulation run utilized deferrable load capacities uniformly distributed in the interval $P_{Ld,k,j,i} \in [100, 6000] W$; inverters with uniformly distributed capacity $P_{inv,k,j}^{max} \in [2000, 5000] W$ and headroom $P_{inv\Delta,k,j,up} \in [0, 2000] W$; and uniformly distributed NLU costs $c_{kji} \in [2, 10]$.

Results shown in Fig. 5 demonstrate that the hierarchical approach for optimal FFR dispatch can achieve the same optimal dispatch result with very similar ($<0.5\%$ difference for $\eta = [15, 20, 25, 30, 35, 40, 45]\%$ and $<1\%$ for all cases) dispatch cost as the optimal MILP solution for each run. Critically, the hierarchical approach can achieve these very nearly optimal dispatch results at a much faster timescale: Fig. 6 shows that the hierarchical approach can be consistently solved within ~ 150 ms with low variance regardless of the problem size, compared to the MILP solution's up to 4-s solution time for the largest problem size (800 transformers or 40,000 total net-load resources). Clearly, the MILP solution's 550–4000+ ms solution time is not fast enough to support FFR dispatch within 500 ms, whereas the designed hierarchical approach provides consistently fast performance to support such FFR dispatch.

VII. CONCLUSIONS

In Part I of this two-part paper, we developed a first-of-its-kind hierarchical control solution that can optimally coordinate tens of thousands of heterogeneous, fast-changing DERs and flexible loads across an electric distribution network to provide FFR to the bulk power system that meets the general FFR service performance requirements. A method for operating solar inverters with the optimal amount of headroom to minimize opportunity cost while ensuring FFR viability of an NLU was also presented. Practical algorithms for achieving optimal FFR dispatch at multiple scales were developed, their optimality proven, and their speed shown to be faster than state-of-the-art approaches. These simulation results provide preliminary validation that coordination of BTM resources in distribution networks can be achieved rapidly enough and on a sufficiently large scale to provide valuable FFR services to the bulk power system. In Part II [32], the entire platform is implemented and experimentally verified on a large-scale PHIL platform with a real-world distribution system model and more than 100 at-power DERs.

REFERENCES

- [1] J. B. Ubertalli and T. B. Littler, "Extending the Reach of Traditional Frequency Control for Fast Responses," in *2019 IEEE Milan PowerTech*, Jun. 2019, pp. 1–6.
- [2] "Frequency Stability Evaluation Criteria for the Synchronous Zone of Continental Europe: Requirements and impacting factors," ENTSO-E, Tech. Rep., Mar. 2016.
- [3] V. Gevorgian, Y. Zhang, and E. Ela, "Investigating the Impacts of Wind Generation Participation in Interconnection Frequency Response," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 3, pp. 1004–1012, Jul. 2015.
- [4] "Inertia: Basic Concepts and Impacts on the ERCOT Grid," Electric Reliability Council of Texas (ERCOT), Tech. Rep., Apr. 2018.
- [5] W. Li, P. Du, and N. Lu, "Design of a New Primary Frequency Control Market for Hosting Frequency Response Reserve Offers From Both Generators and Loads," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 4883–4892, Sep. 2018.
- [6] V. Gevorgian and B. O'Neill, "Advanced Grid-Friendly Controls Demonstration Project for Utility-Scale PV Power Plants," National Renewable Energy Laboratory, Technical Report NREL/TP-5D00-65368, 2016.
- [7] A. J. Arana, J. N. Bank, R. M. Gardner, and Y. Liu, "Estimating speed of frequency disturbance propagation through transmission and distribution systems," in *IEEE Power Systems Conference and Exposition (PSCE)*, 2006.
- [8] A. F. Hoke, M. Shirazi, S. Chakraborty, E. Muljadi, and D. Maksimovic, "Rapid Active Power Control of Photovoltaic Systems for Grid Frequency Support," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 5, no. 3, pp. 1154–1163, Sep. 2017.
- [9] Y.-J. Kim, G. Del-Rosario-Calaf, and L. K. Norford, "Analysis and Experimental Implementation of Grid Frequency Regulation Using Behind-the-Meter Batteries Compensating for Fast Load Demand Variations," *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 484–498, Jan. 2017.
- [10] D. S. Callaway and I. A. Hiskens, "Achieving Controllability of Electric Loads," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 184–199, Jan. 2011.
- [11] A. Molina-Garcia, F. Bouffard, and D. S. Kirschen, "Decentralized Demand-Side Contribution to Primary Frequency Control," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 411–419, Feb. 2011.
- [12] *Annual Energy Outlook 2019*. U.S. Energy Information Administration, Jan. 2019.
- [13] "Electric Storage Participation in Markets Operated by Regional Transmission Organizations and Independent System Operators," U.S. Federal Energy Regulatory Commission, Notice of Proposed Rulemaking, Nov. 2016.
- [14] J. A. Short, D. G. Infield, and L. L. Freris, "Stabilization of Grid Frequency Through Dynamic Demand Control," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1284–1293, Aug. 2007.
- [15] D. J. Hammerstrom, J. Brouss, D. P. Chassin, G. R. Horst, and et al., "Pacific Northwest GridWise™ Testbed Demonstration Projects; Part II. Grid Friendly™ Appliance Project," Pacific Northwest National Laboratory (PNNL), Richland, WA (US), Tech. Rep., 2007.
- [16] H. Karbouj, Z. Rather, D. Flynn, and H. Qzai, "Non-synchronous fast frequency reserves in renewable energy integrated power systems: A critical review," *International Journal of Electrical Power & Energy Systems*, vol. 106, pp. 488–501, 2019.
- [17] J. Lian, J. Hansen, L. D. Marinovici, and K. Kalsi, "Hierarchical decentralized control strategy for demand-side primary frequency response," in *IEEE Power and Energy Society General Meeting*, 2016.
- [18] H. W. Qazi and D. Flynn, "Analysing the Impact of Large-scale Decentralised Demand Side Response on Frequency Stability," *International Journal of Electrical Power and Energy Systems*, Sep. 2016.
- [19] M. R. Vedady Moghadam, R. T. B. Ma, and R. Zhang, "Distributed Frequency Control in Smart Grids via Randomized Demand Response," *IEEE Transactions on Smart Grid*, vol. 5, no. 6, pp. 2798–2809, Nov. 2014.
- [20] Z. Xu, J. Ostergaard, and M. Togeby, "Demand as Frequency Controlled Reserve," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1062–1071, Aug. 2011.
- [21] Y.-Q. Bao, Y. Li, Y.-Y. Hong, and B. Wang, "Design of a hybrid hierarchical demand response control scheme for the frequency control," *IET Generation, Transmission & Distribution*, vol. 9, no. 15, pp. 2303–2310, 2015.
- [22] R. Bhana and T. J. Overbye, "The Commitment of Interruptible Load to Ensure Adequate System Primary Frequency Response," *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 2055–2063, May 2016.
- [23] B. Biegel, L. H. Hansen, P. Andersen, and J. Stoustrup, "Primary control by ON/OFF demand-side devices," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 2061–2071, 2013.
- [24] X. Wu, J. He, Y. Xu, J. Lu, N. Lu, and X. Wang, "Hierarchical Control of Residential HVAC Units for Primary Frequency Regulation," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3844–3856, Jul. 2018.
- [25] A. G. Azar, H. Nazaripouya, B. Khaki, C.-C. Chu, R. Gadh, and R. H. Jacobsen, "A Non-Cooperative Framework for Coordinating a Neighborhood of Distributed Prosumers," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 5, pp. 2523–2534, May 2019.
- [26] Yinliang Xu, Wenxin Liu, and Jun Gong, "Stable Multi-Agent-Based Load Shedding Algorithm for Power Systems," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2006–2014, Nov. 2011.

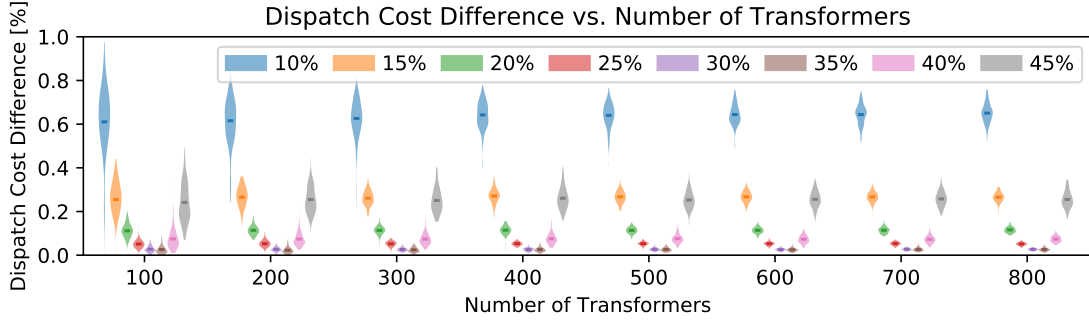


Fig. 5. Full distribution feeder case: Percent difference in dispatch cost achieved by the hierarchical approach versus the MILP optimal solution for varying participation factor (10%–45%).

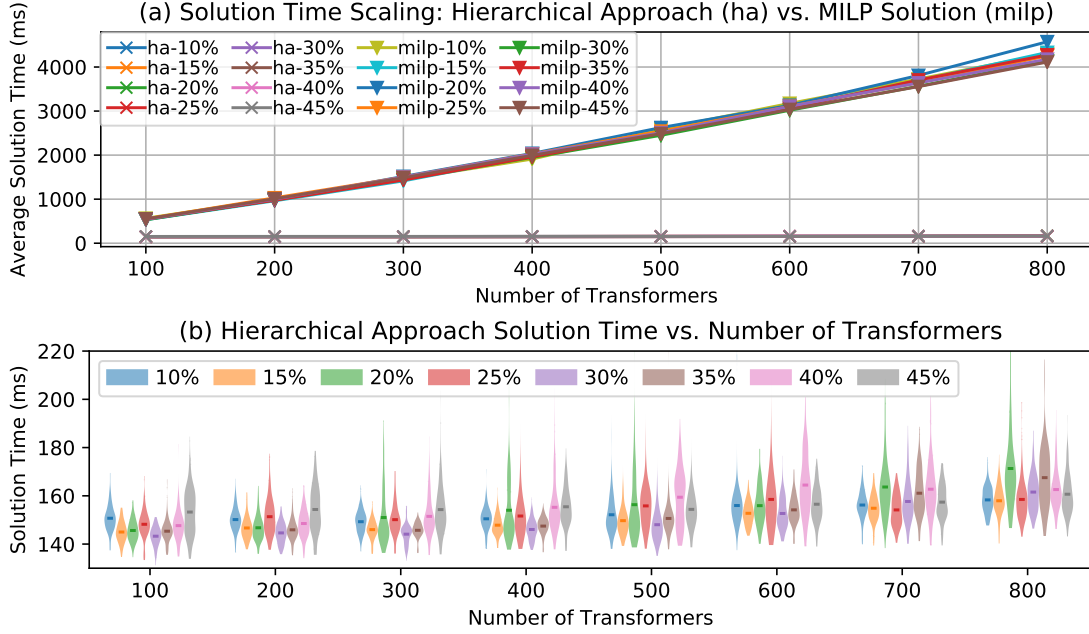


Fig. 6. Full distribution feeder case: Solution time vs. number of transformers (problem size). (a) Comparison of hierarchical approach (“ha”) and MILP solution (“milp”) average solution time based on 155 runs for each case. (b) Violin plot showing distribution of hierarchical approach’s solution times for varying participation factor (10%–45%) for 155 runs each.

- [27] C. Zhao, U. Topcu, and S. H. Low, “Optimal Load Control via Frequency Measurement and Neighborhood Area Communication,” *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3576–3587, Nov. 2013.
- [28] J. Hu, J. Cao, J. M. Guerrero, T. Yong, and J. Yu, “Improving Frequency Stability Based on Distributed Control of Multiple Load Aggregators,” *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1553–1567, Jul. 2017.
- [29] L. Tang and J. McCalley, “Two-Stage Load Control for Severe Under-Frequency Conditions,” *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 1943–1953, May 2016.
- [30] A. M. Prostejovsky, M. Marinelli, M. Rezkalla, M. H. Syed, and E. Guillo-Sansano, “Tuningless Load Frequency Control Through Active Engagement of Distributed Resources,” *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2929–2939, May 2018.
- [31] B. Lundstrom, S. Patel, S. Attree, and M. V. Salapaka, “Fast Primary Frequency Response using Coordinated DER and Flexible Loads: Framework and Residential-scale Demonstration,” in *IEEE Power and Energy Society General Meeting*, Aug. 2018.
- [32] B. Lundstrom, S. Patel, and M. V. Salapaka, “Distribution Feeder-Scale Fast Frequency Response via Optimal Coordination of Net-load Resources—Part II: Large-scale Demonstration,” *IEEE Transactions on Smart Grid*.

APPENDIX A: PROOF (THEOREM 1)

Proof: The net-load system’s actual change in grid power consumption is $\Delta P_G = -\sum_{i=1}^{N_{Ld}} x_i P_{Ld,i} - P_{inv\Delta}$. To ensure problem feasibility, the following must be achieved: $\Delta P_G = \Delta P_G^*$ with $P_{inv\Delta,dn} \leq P_{inv\Delta} \leq P_{inv\Delta,up}$ for $0 \leq P_{inv\Delta,up} \leq (P_{inv}^{max} - P_{inv,base})$ and $0 \leq P_{inv,base} \leq P_{inv}^{max}$.

Consider all possible $-(P_{inv\Delta,up} + \sum_{i=1}^{N_{Ld}} \hat{P}_{Ld,i}) \leq \Delta P_G^* \leq P_{inv\Delta,dn}$ for this problem. If $\Delta P_G^* \leq -\sum_{i=1}^{N_{Ld}} P_{Ld,i}$, a trivial solution of $x_i = 1 \forall i$ and $P_{inv\Delta} = -\sum_{i=1}^{N_{Ld}} x_i P_{Ld,i} - \Delta P_G^*$ is used. Similarly, if $\Delta P_G^* \geq 0$, $x_i = 0 \forall i$ and $P_{inv\Delta} = \Delta P_G^*$. We then consider the remaining cases where $-\sum_{i=1}^{N_{Ld}} P_{Ld,i} < \Delta P_G^* < 0$. For convenience, let $\Delta P^* = -\Delta P_G^*$ and consider the remaining cases where $0 < \Delta P^* < \sum_{i=1}^{N_{Ld}} P_{Ld,i}$.

Let $P_{Ldc} = (P_{Ldc,i})_{i=1}^{N_{Ldc}}$ be the ordered set of all $N_{Ldc} = 2^{N_{Ld}}$ possible load combinations from the elements of P_{Lds} in sequential binary order with $P_{Lds,1}$, the smallest load, being associated with the least significant bit. That is, $P_{Ldc,i} = \sum_{j=1}^{N_{Ld}} y_{ij} P_{Lds,j}$ and $y_{ij} = \begin{cases} 1 & \text{if } (i-1) \bmod 2^j \geq 2^{j-1} \\ 0 & \text{otherwise} \end{cases}$.

Claim 1: For any given $0 < \Delta P^* < \sum_{i=1}^{N_{Ld}} P_{Ld,i}$, there exists at least one $1 \leq j \leq (N_{Ldc} - 1)$, such that $P_{Ldc,j} \leq \Delta P^* \leq P_{Ldc,j+1}$.

Thus, we can consider all cases $0 < \Delta P^* < \sum_{i=1}^{N_{Ld}} P_{Ld,i}$ by examining every discrete interval $P_{Ldc,j} \leq \Delta P^* \leq P_{Ldc,j+1}$ for $1 \leq j \leq (N_{Ldc} - 1)$ and considering problem feasibility and inverter headroom requirements for any ΔP^* in that interval.

Claim 2: Within every discrete interval $P_{Ldc,j} \leq \Delta P^* \leq P_{Ldc,j+1}$, the net-load dispatch problem remains feasible as long as $P_{inv\Delta,up} \geq \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2}$.

Proof: Let $\Delta P^* = P_{Ldc,j} + \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2}$. Within this interval, there are two possible resulting load curtailment outcomes:

Case 1: $\sum_{i=1}^{N_{Ld}} x_i P_{Ld,i} = P_{Ldc,j}$ and so

$$P_{inv\Delta} = \Delta P^* - \sum_{i=1}^{N_{Ld}} x_i P_{Ld,i} = P_{Ldc,j} + \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2} - P_{Ldc,j} = \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2}$$

Case 2: $\sum_{i=1}^{N_{Ld}} x_i P_{Ld,i} = P_{Ldc,j+1}$ and so

$$P_{inv\Delta} = \Delta P^* - \sum_{i=1}^{N_{Ld}} x_i P_{Ld,i} = P_{Ldc,j} + \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2} - P_{Ldc,j+1} = -\frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2}$$

Therefore, $-\frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2} \leq P_{inv\Delta} \leq \frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2}$ and $P_{inv\Delta,up} \geq \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2}$.

We then show that increasing or decreasing $\Delta P^* = P_{Ldc,j} + \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2}$ will not extend these bounds:

Case 1 (decrease): $\Delta P^* = P_{Ldc,j} + \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2} - \epsilon$. Here, $\sum_{i=1}^{N_{Ld}} x_i P_{Ld,i} = P_{Ldc,j}$ because this will minimize $P_{inv\Delta,up}$.

$$\text{Therefore, } P_{inv\Delta} = \Delta P^* - \sum_{i=1}^{N_{Ld}} x_i P_{Ld,i} = P_{Ldc,j} + \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2} - \epsilon - P_{Ldc,j} = \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2} - \epsilon$$

Clearly, $-\frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2} \leq \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2} - \epsilon \leq \frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2}$, so the bounds on $P_{inv\Delta}$ are not extended for this case.

$$\text{Case 2 (increase): } \Delta P^* = P_{Ldc,j} + \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2} + \epsilon.$$

Here, $\sum_{i=1}^{N_{Ld}} x_i P_{Ld,i} = P_{Ldc,j+1}$ because this will minimize $P_{inv\Delta,up}$.

$$\text{Therefore, } P_{inv\Delta} = \Delta P^* - \sum_{i=1}^{N_{Ld}} x_i P_{Ld,i} = P_{Ldc,j} + \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2} + \epsilon - P_{Ldc,j+1} = -\frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2} + \epsilon$$

Clearly, $-\frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2} \leq -\frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2} + \epsilon \leq \frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2}$, so the bounds on $P_{inv\Delta}$ are not extended for this case.

Therefore, increasing or decreasing ΔP^* does not extend the bounds on $P_{inv\Delta}$, so $P_{inv\Delta,up} \geq \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2}$ remains.

Next, we show that the bounds $-\frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2} \leq P_{inv\Delta} \leq \frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2}$ on $P_{inv\Delta}$ cannot be made narrower (e.g., $-\frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2} + \epsilon \leq P_{inv\Delta} \leq \frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2} - \epsilon$) without causing problem infeasibility. To show this, choose $\Delta P^* = P_{Ldc,j} + \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2}$. As before, there are then two possible cases:

Case 1: $\sum_{i=1}^{N_{Ld}} x_i P_{Ld,i} = P_{Ldc,j}$ and so

$$P_{inv\Delta} = \Delta P^* - \sum_{i=1}^{N_{Ld}} x_i P_{Ld,i} = P_{Ldc,j} + \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2} - P_{Ldc,j} = \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2}$$

However, $\frac{P_{Ldc,j+1}-P_{Ldc,j}}{2} \not\leq \frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2} - \epsilon$ and therefore the problem is infeasible with bounds $-\frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2} + \epsilon \leq P_{inv\Delta} \leq \frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2} - \epsilon$.

Case 2: $\sum_{i=1}^{N_{Ld}} x_i P_{Ld,i} = P_{Ldc,j+1}$, so:

$$P_{inv\Delta} = \Delta P^* - \sum_{i=1}^{N_{Ld}} x_i P_{Ld,i} = P_{Ldc,j} + \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2} - P_{Ldc,j+1} = -\frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2}$$

However, $-\frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2} \not\geq -\frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2} + \epsilon$ and therefore the problem is infeasible with bounds $-\frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2} + \epsilon \leq P_{inv\Delta} \leq \frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2} - \epsilon$.

Thus, we have shown that the bounds of at least $-\frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2} \leq P_{inv\Delta} \leq \frac{(P_{Ldc,j+1}-P_{Ldc,j})}{2}$ on $P_{inv\Delta}$ are required to maintain problem feasibility, and therefore $P_{inv\Delta,up} \geq \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2}$.

Claim 3: The maximum $P_{inv\Delta,up,j} \geq \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2}$ across all discrete intervals $P_{Ldc,j} \leq \Delta P^* \leq P_{Ldc,j+1}$ for all $1 \leq j \leq (N_{Ldc} - 1)$, and therefore for any $0 < \Delta P^* < \sum_{i=1}^{N_{Ld}} P_{Ld,i}$, can be obtained using the simplified expression:

$$P_{inv\Delta,up}^{max} = \alpha = \max_{1 \leq i \leq N_{Ld}} \frac{P_{Lds,i} - \sum_{n=1}^{i-1} P_{Lds,n}}{2} \quad (19)$$

Proof: $P_{inv\Delta,up}^{max}$ can be obtained by simply examining the $P_{inv\Delta,up,j} \geq \frac{P_{Ldc,j+1}-P_{Ldc,j}}{2}$ at each discrete interval $j = 1 \dots N_{Ldc} - 1$ and taking the maximum. That is:

$$P_{inv\Delta,up}^{max} = \max_{1 \leq j \leq N_{Ldc} - 1} \frac{P_{Ldc,j+1} - P_{Ldc,j}}{2} \quad (20)$$

Note that each $\frac{P_{Ldc,j+1}-P_{Ldc,j}}{2}$ can be evaluated by considering the constituent loads $P_{Lds,i}$ for all $i = 1 \dots N_{Ld}$ that change state (i.e., $y_{ij} \neq y_{i(j+1)}$) between $P_{Ldc,j}$ and $P_{Ldc,j+1}$. Because P_{Ldc} is in sequential binary order and P_{Lds} is in sequential order, the sequence of loads changing as j increases $j = 1 \dots N_{Ldc} - 1$ follows a binary pattern wherein:

for each i in $i = 1 \dots N_{Ld}$:

$$\begin{aligned} (P_{Ldc,k+1} - P_{Ldc,k}) &= P_{Lds,i} \\ \text{where } k &= 2^i \cdot m - 2^{i-1} \\ \text{for all } m &= 1 \dots 2^{N_{Ld}-1} \end{aligned}$$

This repeating pattern has the consequence that the only unique $(P_{Ldc,j+1} - P_{Ldc,j})$ differences not already considered in earlier intervals $(P_{Ldc,j+1} - \epsilon - P_{Ldc,j} - \epsilon)$ (where $\epsilon \in \mathbb{Z}$, $\epsilon = 1 \dots (j-1)$) occur at the intervals $(P_{Ldc,(2^i-2^{i-1}+1)} - P_{Ldc,(2^i-2^{i-1})}) \forall i = 1 \dots N_{Ld}$ where:

$$(P_{Ldc,(2^i-2^{i-1}+1)} - P_{Ldc,(2^i-2^{i-1})}) = P_{Lds,i} - \sum_{n=1}^{i-1} P_{Lds,n} \quad \forall i = 1 \dots N_{Ld}$$

Therefore:

$$\begin{aligned} P_{inv\Delta,up}^{max} &= \max_{1 \leq j \leq (N_{Ldc} - 1)} \frac{P_{Ldc,j+1} - P_{Ldc,j}}{2} \\ &\Rightarrow P_{inv\Delta,up}^{max} = \max_{1 \leq i \leq N_{Ld}} \frac{(P_{Ldc,(2^i-2^{i-1}+1)} - P_{Ldc,(2^i-2^{i-1})})}{2} \\ &= \max_{1 \leq i \leq N_{Ld}} \frac{P_{Lds,i} - \sum_{n=1}^{i-1} P_{Lds,n}}{2} \end{aligned}$$

APPENDIX B: PROOF (THEOREM 2)

Proof: When solving the optimization problem (5)–(8), a final optimal solution $(x^\mu, P_{inv\Delta}^\mu)$ is obtained. Note that $(x^\mu, P_{inv\Delta}^\mu) \in \Omega^\mu$. When solving the optimization problem (9)–(12), x^ν is obtained, and then $P_{inv\Delta}^\nu$ is set as given in (13). This yields a final solution $(x^\nu, P_{inv\Delta}^\nu)$. Note that $(x^\nu, P_{inv\Delta}^\nu) \in \Omega^\nu$.

First, we show $\mu \leq \nu$:

We will compare $(x^\nu, P_{inv\Delta}^\nu) \in \Omega^\nu$ of the ν problem and the feasible set Ω^μ of the μ problem. To begin, consider the result of (13):

$$P_{inv\Delta}^\nu = \sum_{i=1}^{N_{Ld}} (1 - x_i^\nu) P_{Ld,i} - P_{inv,base} - P_G^*$$

By (10), $\sum_{i=1}^{N_{Ld}} (1 - x_i) P_{Ld,i} - P_{inv,base} - P_G^* \leq P_{inv\Delta,up}$ so:

$$P_{inv\Delta}^\nu \leq P_{inv\Delta,up}$$

Equation (11) guarantees $P_{inv\Delta}^\nu \geq P_{inv\Delta,dn}$, and therefore:

$$P_{inv\Delta,dn} \leq P_{inv\Delta}^\nu \leq P_{inv\Delta,up}$$

This and (13) satisfy (6)–(7), therefore $P_{inv\Delta}^\nu$ meets all constraints of Ω^μ placed on $P_{inv\Delta}$.

By (12), $x^\nu \in \{0, 1\}$, which satisfies (8), so x^ν meets all constraints of Ω^μ placed on x . Therefore:

$$(x^\nu, P_{inv\Delta}^\nu) \in \Omega^\mu \quad (21)$$

As:

$$\mu = \min_x \sum_{i=1}^{N_{Ld}} x_i P_{Ld,i}$$

$$\text{subject to } (x^\mu, P_{inv\Delta}^\mu) \in \Omega^\mu$$

Because $(x^\nu, P_{inv\Delta}^\nu) \in \Omega^\mu$, there exists:

$$\nu = \sum_{i=1}^{N_{Ld}} x_i^\nu P_{Ld,i}$$

and because μ is the minimum $\sum_{i=1}^{N_{Ld}} x_i P_{Ld,i}$ for any $x \in \Omega^\mu$, therefore:

$$\mu \leq \nu \quad (22)$$

We now show $\nu \leq \mu$:

Examining $(x^\mu, P_{inv\Delta}^\mu) \in \Omega^\mu$ and comparing Ω^μ to the constraint set Ω^ν defined by (10)–(12):

Combining (6) with $P_{inv\Delta} \leq P_{inv\Delta,up}$ yields:

$$\sum_{i=1}^{N_{Ld}} (1 - x_i) P_{Ld,i} - P_{inv,base} - P_{inv\Delta,up} \leq P_G^*$$

which matches (10). Rearranging (6) yields:

$$\sum_{i=1}^{N_{Ld}} (1 - x_i) P_{Ld,i} - P_{inv,base} - P_{inv\Delta} = P_G^*$$

which satisfies (11). Finally, (8) satisfies (12) directly. All constraints of Ω^ν are satisfied, therefore, $(x_i^\mu, P_{inv\Delta}^\mu) \in \Omega^\nu$.

Then, from (9)–(13):

$$\nu = \min_x \sum_{i=1}^{N_{Ld}} x_i P_{Ld,i}$$

$$\text{subject to } (x^\nu, P_{inv\Delta}^\nu) \in \Omega^\nu$$

Because $(x^\mu, P_{inv\Delta}^\mu) \in \Omega^\nu$, there exists:

$$\mu = \sum_{i=1}^{N_{Ld}} x_i^\mu P_{Ld,i}$$

and because ν is the minimum $\sum_{i=1}^{N_{Ld}} x_i P_{Ld,i}$ for any $x \in \Omega^\nu$:

$$\nu \leq \mu \quad (23)$$

Finally, by (22) and (23):

$$\mu = \nu$$



Blake Lundstrom (M'11–SM'19) received the B.S. and M.S. degrees in electrical engineering from the Colorado School of Mines, Golden, CO, USA in 2010 and 2013, respectively, and the Ph.D. degree in electrical engineering from the University of Minnesota, Minneapolis, MN, USA in 2020. He is a Senior Research Engineer in the Power System Engineering Center at the National Renewable Energy Laboratory, Golden, CO, USA, where he has been employed since 2009. His research involves design of control techniques for managing large-scale deployments of distributed energy resources, distribution system modeling, power electronics controls development, and the use and development of real-time hardware-in-the-loop techniques for energy systems integration research. Dr. Lundstrom is a Vice Chair for IEEE P2004: *Recommended Practice for Hardware-in-the-Loop Simulation Based Testing of Electric Power Apparatus and Controls* and is a licensed Professional Engineer in the State of Colorado.



Sourav Patel (S'18) received his B.Tech. degree in Instrumentation and Control Engineering from the National Institute of Technology, Jalandhar, India, in 2011. In the same year, he joined National Thermal Power Corporation Ltd. in Kaniha, Odisha, India as a Control and Instrumentation Engineer. He received his M.S. degree in Electrical Engineering from the University of Minnesota, Minneapolis, MN, USA, in 2018 where currently, he is working towards his Ph.D. degree in Electrical Engineering. His research interests include control and systems theory, coordination and communication protocols for Distributed Energy Resources towards smart grid applications.



Murti V. Salapaka (F'19) received the B.Tech. degree in Mechanical Engineering from the Indian Institute of Technology, Madras, in 1991 and the M.S. and Ph.D. degrees in Mechanical Engineering from the University of California at Santa Barbara, in 1993 and 1997, respectively. He was a faculty member in the Electrical and Computer Engineering Department, Iowa State University, Ames, IA, USA, from 1997 to 2007. Currently, he is the Director of Graduate Studies and the Vincentine Hermes Luh Chair Professor in the Electrical and Computer Engineering Department, University of Minnesota, Minneapolis, MN, USA. His research interests include control and network science, nanoscience and single molecule physics. Dr. Salapaka received the 1997 National Science Foundation CAREER Award and is an IEEE fellow.