Real-Time Grid and DER Co-Simulation Platform for Testing Large-Scale DER Coordination Schemes

Adil Khurram[®], *Member, IEEE*, Mahraz Amini[®], *Member, IEEE*, Luis A. Duffaut Espinosa[®], *Senior Member, IEEE*, Paul D. H. Hines[®], *Senior Member, IEEE*, and Mads R. Almassalkhi[®], *Senior Member, IEEE*

Abstract-Distributed energy resources (DERs), such as responsive loads and energy storage systems, can help grid operators to more effectively balance supply and demand and manage network constraints. However, consumer acceptance of load coordination schemes depends on ensuring customer quality of service (QoS), thus requiring the careful management of device-level constraints. Simultaneously managing both devicelevel and system-level objectives is a difficult problem. Most potential solutions to the DER coordination problem are in some way distributed, with decision logic running both at the device (such as a thermostat) and the coordinator, with potentially complex communications between devices and one or more coordinators. It is difficult to effectively test distributed DER control schemes, particularly when there are thousands of devices with both device-level and grid-level constraints and asynchronous communications. This paper addresses this challenge by presenting a novel real-time grid-and-DER cyber-physical co-simulation platform for testing advanced DER coordination schemes and characterizing the behaviors of a DER fleet. The co-simulation platform is particularly suitable for: i) testing the real-time performance of a large fleet of DERs delivering advanced grid services; ii) characterizing the distribution of states for a large fleet of managed DERs; and iii) incorporating practical limitations of DERs and communications and analyzing the effects on fleet-wide performance. To illustrate these benefits, we demonstrate how the platform can be used to test a fleet with thousands of DERs coordinated using Packetized Energy Management (PEM) and interacting with a realistic transmission system model.

Index Terms—Cyber-physical systems, distributed energy resources, demand dispatch, packetized energy management.

Manuscript received 7 April 2021; revised 29 September 2021; accepted 26 April 2022. Date of publication 20 June 2022; date of current version 21 October 2022. This work was supported by the U.S. Department of Energy's Advanced Research Projects Agency-Energy (ARPA-E) under Award DE-AR0000694. The work of Mads R. Almassalkhi was supported in part by the National Science Foundation (NSF) under Award ECCS-2047306. Paper no. TSG-00528-2021. (Corresponding author: Adil Khurram.)

Adil Khurram is with the Department of Mechanical and Aerospace Engineering, University of California at San Diego, San Diego, CA 92093 USA (e-mail: akhurram@ucsd.edu).

Mahraz Amini is with the Market Fundamentals Department, National Grid, Waltham, MA 02451 USA (e-mail: mahraz.amini@nationalgrid.com).

Luis A. Duffaut Espinosa and Mads R. Almassalkhi are with the Department of Electrical and Biomedical Engineering, University of Vermont, Burlington, VT 05405 USA (e-mail: lduffaut@uvm.edu; malmassa@uvm.edu).

Paul D. H. Hines was with the Department of Electrical and Biomedical Engineering, University of Vermont, Burlington, VT 05405 USA. He is now with EnergyHub, Brooklyn, NY 11217 USA (email: hines@energyhub.net).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSG.2022.3184491.

Digital Object Identifier 10.1109/TSG.2022.3184491

NOMENCLATURE

SoC of i-th DER.

\underline{x} , \overline{x}	Lower and upper SoC of individual DER.
x_{set}	Dead-band setpoint of individual DER.
ζ	Operating state of DER (charge, discharge or
	off).
P_{ζ}	Rated power.
L	Tank size of electric water heaters.
c	Specific heat constant.
ρ	Density of water at 25° C.
\mathcal{C}	Capacity of energy storage systems.
η_{ζ}	Charging/discharging efficiency of DER.
Δt	Simulation time-step.
z	Aggregate SoC of DER fleet.
δ_{ζ}	Packet length.
m_R	Mean time to request.
q	Probability mass function in macromodel.
β	Proportion of accepted packets in PEM.
eta^-	Proportion of expiring packets in PEM.
P_{nom}	Nominal power demand of the fleet.
P_{ref}	Reference power.
P_{dem}	Power demand of the fleet.
$n_{r,\mathrm{acc}}^{\mathrm{c}}, n_{r,\mathrm{acc}}^{\mathrm{d}}$	Number of accepted charging and discharging
	requests.
$n_r^{\rm c}, n_r^{\rm d}$	Number of charging and discharging requests
	made.
R	Droop parameter.
B, k	AGC gain parameters.

I. INTRODUCTION

Generator schedules.

THE DRIVE to reduce greenhouse gas emissions and declining costs for distributed renewable generation are leading to rapid increases in wind and solar generation. However, despite their low emissions profile, wind and solar power generation vary rapidly in time, motivating the need for additional balancing resources. While it has long been argued that distributed resources can assist with balancing [1]–[4], the availability of Internet connected and controllable distributed energy resources (DERs) make it possible to coordinate DERs for increasingly sophisticated grid services. DERs such as electric water heaters (EWHs) and energy storage systems (ESS)

 $P_{\rm gen}^{\rm sch}$

Platform	Asynch. comm.	Comm. net- work sim.	Coordinator controller design	Scalable	Real-time grid sim.	Heterogeneous, diverse	Asynch. DER sim.	PHIL/CHIL	Other tools
UVM	√	×	√	√	√	√	√	×	×
NREL	×	✓	×	√	✓	√	√	✓	✓
Sandia DETL	×	✓	×	×	✓	×	✓	✓	✓
PNWSGD, TESP	×	✓	×	×	×	✓	✓	✓	✓
SGRS CMU	×	×	✓	✓	×	X	×	×	×
Vanderbilt	√	✓	√	×	×	×	×	×	×
Power Matcher	✓	×	✓	×	×	✓	×	✓	×
UCR	×	×	√	×	√	×	×	√	×

TABLE I
COMPARISON OF REAL-TIME SIMULATION PLATFORMS

can be coordinated to provide grid services without degradation in quality of service (QoS) for the customer [4]–[12]. The capability to aggregate and control DERs to provide services to the grid with QoS guarantees is called demand dispatch.

The main ideas underlying modern demand dispatch have existed for decades [2]-[4]. The pioneering work focusing on direct load control of electric water heaters [4] provided analytical tools to evaluate the effect of control strategies such as cold load pick-up on a group of EWHs. Since then, several DER coordination schemes have been proposed and validated in simulations. Some of the schemes that are relevant to this work are [6]-[12]. These schemes either assume full control over DER's operating state and knowledge of its state of charge (SoC) in a centralized coordinator [5] or influence power consumption of DERs indirectly by transmitting a control command at regular intervals in a distributed coordinator. Packetized energy management (PEM) is a distributed load coordination scheme in which DERs request the PEM coordinator to consume power which can be accepted or rejected as required [9]-[12]. Distributed control schemes, like PEM, assume access to limited DER information but employ modelbased state estimation schemes to estimate the SoC of the fleet. Nonetheless, all modern demand dispatch methods require some level of communications between DERs and coordinators, making these systems good examples of cyber-physical systems (CPS).

Recent advances in sensing and connectivity have led to widespread adoption of cyber-physical systems as a framework in which to incorporate, model, and analyze the interactions between the cyber (e.g., communication networks) and the physical (e.g., grid) [13]–[16]. In this regard, CPS test-beds have been developed within the smart grid area that consist of a combination of physical and simulated components [17]–[26]. Real-world DERs and controllers are connected to

grid simulators along with a combination of diverse and heterogeneous simulated DERs constituting power and controller hardware in the loop simulations (P-HIL and C-HIL, respectively). DERs typically communicate using standard protocols over the Internet or through private communications channels. Furthermore, software used by transmission and distribution system operators, such as advanced distribution management systems (ADMS) or distributed energy resource management systems (DERMS), can be used to manage both simulated and real DERs. CPS test-beds [17]–[26] generally aim to incorporate one or several of these objectives.

Existing CPS testing platforms range from demonstration projects consisting of real-world DERs only, such as Pacific Northwest Smart Grid Demonstration (PNWSGD) [24] project, to a combination of simulated and real devices. National Renewable Energy Laboratory's DER test facility (DERTF) [17] and Sandia National Lab's distributed energy technology lab (DETL) [21] are among the most comprehensive existing CPS test-beds. These systems incorporate real-time grid simulation, P-HIL and C-HIL, simulated and real-world DERs and ADMS. DERTF and DETL are geared towards utilities and vendors and aim to provide a platform for rapid control prototyping (e.g., islanding controllers or PV inverter systems), protection and automation, testing of ADMS, standards development etc., under a wide range of simulated and emulated scenarios. PowerMatcher is another tool that has been used for simulation and deployment of DERs in the real-world but on a small scale only. In a more academic setting, a P-HIL test-bed has been developed for grid-connected battery energy storage systems at University of California, Riverside (UCR) [22]. Other platforms [25], [27] focus on testing of advanced DER coordination schemes in a scalable but not real-time environment that incorporates grid communication. Table I compares some key characteristics of these CPS platforms. Also other tools in Table I correspond to softwares such as ADMS.

Most of the more complex testing systems are specially designed for transient stability analysis and cyber-security performance assessment, as shown in Table I. As a result, running these simulations requires detailed models of both how the DERs interact with the grid, and also how they communicate with each other and with coordinators. As such, several of the test systems simulate complex inter-device communications systems. Similarly, many include detailed transient stability simulation systems that allow for one to study the response of systems to faults, measurement noise, communication latency, and malicious data injection. However, due to the complexity of accurately simulating both inter-device communications and power system dynamics, most existing CPS test systems struggle to simulate more than a few hundred devices.

Load coordination schemes, on the other hand, require thousands of DERs to create a net-effect reasonable for providing grid services. Aggregate behavior of DERs under these schemes has been extensively studied and subsequently modeled in [6]–[12]. Although, physical limitations such as communication delays have been incorporated in synchronized MATLAB based simulated environments [28], however, the communication between DERs and the coordinator occurs asynchronously. This is because DERs use a local clock for operation that introduces practical limitations associated with device operation and communication. It is not possible to identify and then model such limitations in MATLAB simulations. Furthermore, the CPS test-beds mentioned above either ignore these effects or consider them on a small scale limited to fewer number of DERs than needed for demand dispatch.

As indicated in Table I there is a need for CPS testing systems that can simulate thousands of diverse and heterogeneous DERs with asynchronous communications and real-time interactions with power system dynamics. To address this gap, this paper presents a co-simulation platform developed at University of Vermont (UVM), with the following unique contributions.

- The proposed platform combines grid simulation and asynchronous simulation of diverse DER fleets in a real-time co-simulation environment.
- The co-simulator is agnostic to the type of DER coordination scheme.
- The communication mechanisms between DERs and the coordinator very closely align with those used by IoT-enabled devices in real-world demand dispatch. Furthermore, communication latencies can be incorporated to analyze their effect on fleet-wide performance.
- DER simulation is not only agnostic to the operating system (e.g., Linux and Windows), but can also make use of embedded micro-controllers, making the platform scalable to thousands of DERs.
- Advanced demand dispatch models are incorporated in the DER coordinator that enable online state estimation and testing real-time performance of large fleets of DERs in delivering grid services, including frequency regulation. Relevant case studies are presented for a distributed

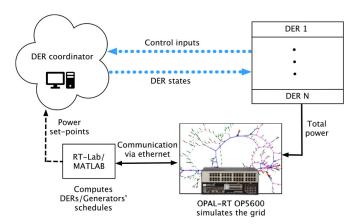


Fig. 1. Cyber-physical platform overview: The transmission grid is simulated on OP5600 and AGC based dispatch is realized on a host PC that dispatches demand. The packetized load is emulated on a high performance PC that is aggregated by a python based server.

demand dispatch scheme called packetized energy management. Aggregate models of PEM that have previously been tested in MATLAB simulation are now validated on the real-time co-simulator.

It should be noted here that communication protocols often used for communications with utility assets, such as DNP3, are not implemented since the focus of this platform is on (behind-the-meter) demand dispatch, rather than distribution automation. Illustrative results are presented for packetized energy management, a bi-directional DER-to-Coordinator demand dispatch scheme [9]–[12], for which preliminary validation is carried out in [29] along with the early field-trial testing with customer-owned electric water heaters in [30].

The remainder of the paper is organized as follows. In Section II, the grid-and-DER co-simulation platform is presented. Section III describes packetized energy management of DERs. Sections IV–VI subsequently deal with the implementation of advanced aggregate models for control and estimation, validation of grid services and identification of practical limitations associated with demand dispatch. Finally, Section VII concludes the paper.

II. REAL-TIME GRID-AND-DER CO-SIMULATION PLATFORM

This section describes the real-time cyber enabled platform developed in this work. Real-time herein refers to the orders of tens of milliseconds. The proposed validation platform consists of four components, (i) grid simulator, (ii) DER emulator, (iii) communication channel, (iv) DER coordinator or aggregator. The over-arching goal of this platform is to provide means to study and demonstrate the effects of large scale demand dispatch on the grid in a realistic setting as well as identify potential DER related challenges associated with such deployment.

A. Grid Simulator

The electric grid is simulated using the OP5600 real-time grid simulator and ePHASORSIM from OPAL-RT. The grid is

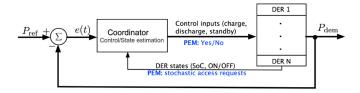


Fig. 2. Closed-loop feedback system for demand dispatch to track P_{ref} based on the aggregate net-load measured by the coordinator and control inputs to DERs.

modeled with a standard equivalent positive-sequence singlephase constant-power AC model in ePHASORSIM. In this paper, the electric grid is based on the Vermont Electric Power Company's (VELCO) transmission system and its details are further discussed in Section V-A1.

B. DER Coordinator

The coordinator managing DERs sends control signals to the fleet depending upon the power reference $P_{\rm ref}$ provided by the utility and or aggregator, the total power consumption of the fleet $P_{\rm dem}$ and measurements regarding DER's states. The block diagram in Fig. 2 represents a general DER coordination scheme where the control signal can be ON/OFF commands for individual DERs or an ON/OFF switching fraction that is broadcast to the entire fleet [8]. In case of PEM [9], it is simply a YES/NO response to a DER's request which is described in detail in the next section.

The DER coordinator is implemented in Python using requests module and an open-source event-driven networking engine called Twisted. Specifically, a TCP server listens for Hypertext Transfer Protocol (HTTP) messages from the emulated DERs. HTTP is a standard protocol for asynchronous communication between Web servers and clients and is highly inter operable over several programming languages (C/C++ or python) and platforms (Linux or micro-controllers). Therefore, the coordinator leverages HTTP GET and POST methods to realize DER-coordinator communication for transmitting states and control inputs. Different demand dispatch schemes can be implemented in this setup, such as but not limited to, the load control of [8] or the PEM based demand dispatch [9]. Only the nature of control inputs and state information shared with the coordinator needs to be configured.

Finally, the asynchronous mechanism used by the HTTP-based simulation and incorporated in this work very closely represents the cloud-based implementation of demand dispatch used in many cloud-based DERMS platforms. The only element missing is security layer (i.e., HTTPS with TLS 1.2 or similar). IEEE 2030.5, for example, is typically implemented using HTTPS.

C. Emulation of DERs

DER logic is written in C/C++ based on a simplified first order state of charge dynamics. The parameter of DERs, such as rated power, tank size, end-use consumption etc., can be different for each DER which is set at the start of the simulation. Furthermore, since DER simulator is asynchronous and

allocated its own thread, the platform can consider *any* DER discrete-time representation, such as high-order nonlinear, linear, or even a performance map. That is, a first-order DER model was mainly out of convenience.

A DER is initialized as a thread that executes its logic as a background process in the processor of a high performance computer. Using this setup, thousands of DERs can be emulated by simply spawning a thread for each DER in the processor's memory. Operating system used for simulations in this paper is Linux that manages the internal scheduling of the threads. Each thread then emulates a DER equipped with its own clock. The key difference between this implementation and other MATLAB based simulations is that each DER executes its logic independently of other DERs. To enable demand dispatch, communication between the DERs and the DER coordinator is achieved over the Internet via the HTTP requests mechanism. Demand dispatch via Internet enabled DERs is becoming more common, such as those presented in [30], which is accurately represented in this platform. Although, the results presented in this paper are for PEM, however, the DER implementation is flexible enough to allow any type of demand dispatch.

D. Communication Channel

As mentioned earlier, the communication between DERs and the coordinator occurs by means of HTTP requests over the Internet. This type of implementation allows one to identify practical limitations associated with real-world deployment that includes but not limited to communication latency, packet loss and loss of Internet connection as highlighted in the next remark. The effect of communication latency is specifically studied in Section VI.

Remark 1: Both the DER emulator and the coordinator are executed on the same PC, however, it is important to highlight that the coordinator and each of the emulated DER are executed as a separate thread in the computer's memory. In short this means that decisions made by the coordinator will not affect the requesting DERs instantaneously rather control inputs to DERs must pass through the communication channel that incorporates delays as well as dynamics of the channel itself. This is essential to ensure asynchronism is maintained which has a significant impact in real-world DERs but usually ignored in MATLAB based simulation studies.

In the proposed platform, the total power of the fleet can be communicated to the grid simulator by two methods. In the first method, a micro-controller ESP8266 reads $P_{\rm dem}$ via the HTTP GET/POST method from the coordinator, converts kWs into an equivalent analog voltage between $0-10{\rm V}$ which is then connected to one of the OPAL-RT's analog I/O pins. This method was demonstrated in [29]. However, the analog signal is prune to errors due to kW to Volt conversion in the micro-controller and then again from Volts to kW in OPAL-RT. Therefore, the second method adopted in this paper that makes use of the HTTP GET/POST methods to read $P_{\rm dem}$ directly from the coordinator.

III. PACKETIZED ENERGY MANAGEMENT PRELIMINARIES

In order to test and demonstrate the capabilities of this CPS testing platform, we used it to test a demand dispatch method known as packetized energy management (PEM). PEM coordinates fleets of DERs through its unique request-response mechanism. DERs asynchronously request to consume power from the grid or inject power into the grid which begets an energy packet. The platform developed herein captures PEM's asynchronous communication mechanism by the thread-based emulation of DERs in real-time.

A. Device-Driven Coordination of DERs

A DER with SoC x, operates within a deadband $[x, \bar{x}]$ around a user-specified setpoint x_{set} where \underline{x} and \overline{x} are the lower and upper bounds of the deadband. Then, based on their SoC, PEM enabled DERs can be in one of the four logical modes, i.e., (i) charge, (ii) discharge, (iii) standby and (iv) optout [9]. In the standby mode with $x \in [x, \overline{x}]$, DERs requests the aggregator to either consume power in the charge mode from the grid or inject power into the grid in discharge mode which is called a packet request. If a charging (discharging) packet request is accepted, then the DER transitions to charge (discharge) mode for a specified period of time called charging (discharging) packet length and is denoted as δ_1 (δ_{-1}). Furthermore, the opt-out mode is included in PEM that ensures QoS by allowing DERs to temporarily opt-out of PEM if x < xand charge until SoC has sufficiently recovered. DERs can also opt-out due to excessive energy $(x > \bar{x})$. The request mechanism is described next.

The probability that the i-th DER in "standby" mode with SoC $x_i[k]$ at time k requests a charging packet is designed to be the following cumulative exponential distribution function,

$$g_{\mu}(x_i[k]) = 1 - e^{-\mu(x_i[k])\Delta t}$$
 (1)

$$g_{\mu}(x_{i}[k]) = 1 - e^{-\mu(x_{i}[k])\Delta t}, \qquad (1)$$

$$\mu(x_{i}[k]) = \begin{cases} 0, & \text{if } x[k] \geq \overline{x} \\ m_{R}\left(\frac{\overline{x} - x_{i}[k]}{x_{i}[k] - \underline{x}}\right)\left(\frac{x_{\text{set}} - \underline{x}[k]}{\overline{x}[k] - x_{\text{set}}}\right), & \text{if } x[k] \in (\underline{x}, \overline{x}) \\ \infty, & \text{if } x[k] \leq \underline{x} \end{cases} \qquad (2)$$

where, $\mu(x_i[k])$ is called the rate parameter, m_R is a design parameter called the mean time to request. The discharge probability of request can be obtained in a similar manner [9]. The probability of request function (g_{μ}) is designed to ensure that a DER with lower SoC has a greater probability of making a request as compared to the one with higher energy.

The SoC evolves according to the first order linear difference equation,

$$x_i[k+1] = x_i[k] + \Delta t(\phi_{\zeta} P_{\zeta} \zeta[k] + \phi_{s}[k] x_i[k] + \phi_{u}[k]),$$
 (3)

where $\zeta \in \{-1, 0, 1\}$ is the state of the thermodynamic switch, ϕ_{ζ} is the charging or discharging efficiency, ϕ_{s} represents the standing losses and ϕ_u models the effect of end-use consumption. When $\zeta = 1$ ($\zeta = -1$) then the DER is charging (discharging) at the power transfer rate P_1 (P_{-1})kW, efficiency η_1 (η_{-1}) and the DER is in standby mode when $\zeta = 0$. It should be noted here that the platform is not limited to only linear DER models which can be replaced by complex higher order models. Furthermore, the choice of linear model is leveraged in the next subsection to model aggregate behavior of DER fleets. In these models, it is reasonable for the coordinator to view arbitrary non-linear dynamics as a linear system to simplify models for control design. Otherwise significant effort in terms of modeling and controller design is required that are difficult to analyze and implement in a real-time setting.

Case studies presented in this paper illustrate the capabilities of the platform using thermostatically controlled loads and energy storage system (ESS). However, any other flexible DER and load can also be implemented on this platform by appropriately modifying (3). For TCLs such as electric water heaters that can only charge, that is $\zeta \in \{0, 1\}$, let $\Theta_{\text{EWH}} = \{P_{\zeta}, L, x_{\text{set}}, [\underline{x}, \overline{x}]\}$ be the set of parameters. Then in (3), $\phi_{\zeta} = (\eta_{\zeta} P_{\zeta})(c\rho L)^{-1}$ where η_{ζ} is the efficiency, $c = 4.186 \text{ (kJ)}(\text{kg}^{\circ}C)^{-1}$ is the specific heat constant, $\rho = 0.990 \text{ kgL}^{-1}$ is the density of water when close to 50 °C and L is the tank size in liters. Furthermore, $\phi_s[k] = (x[k] (x_a)(x[k]\tau)^{-1}$ and $\phi_u[k] = Q[k](c\rho L)^{-1}$ where, $x_a = 21^{\circ}C$ is the ambient temperature, $\tau = 150 \times 3600$ seconds is the standing loss time constant to ambient temperature and O[k] is the heat loss from the tank due to customer driven water usage and is modeled as Poisson random pulses [11]. Similarly, energy storage systems (ESSs) with capacity C in kWh can both charge and discharge with efficiencies $\phi_1 = \eta_1$ and $\phi_{-1} = \eta_{-1}$ respectively. End-use consumption and standing losses for ESSs are assumed to be negligible, $\phi_s = \phi_u = 0$ resulting in the parameter set $\Theta_{ESS} = \{P_{\zeta}, C, \eta_{\zeta}, x_{set}, [\underline{x}, \overline{x}]\}.$

The dynamics of an individual DER is modeled using (3) and employed for emulation in the platform. However, individual SoC is not generally measured in distributed coordinators. Controller design such as the one in [12] require some information regarding the current state of the fleet and it is not feasible to use (3) at the coordinator for this purpose due to scalability issues. This means that the number of states increases with the fleet size and becomes intractable for thousands of DERs. Population based models, on the other hand, can capture the aggregate behavior of the fleet using a fewer number of states and are specifically suitable for control design at the coordinator [11] as described next.

B. Macromodel for Aggregate Population

State bin transition model, referred to as macromodel in this paper, is a Markovian model of DERs. Macromodel captures the average behaviour of a sufficiently large fleet according to the law of large numbers which can then be used at the coordinator for control design. It is shown in [12] that as the number of DERs increases the modeling error in the macromodel decreases. Furthermore, the aggregate dynamics of a fleet consisting of 250 DERs is sufficiently captured by the macromodel with normalized modeling error less than 2.5%. The macromodel is presented next.

To obtain the macromodel, the continuous interval $[x, \bar{x}]$ is divided into a set of discrete states or bins. Each of the operating modes in PEM then corresponds to identical copies of these binned states. For example, in the case of ESSs, four sets of bin states model the charge, discharge, standby and opt-out modes. We then keep track of the proportion of DERs in each bin in the vector q[k] which is called as the probability mass function (PMF). The dimension of q[k] corresponds to the total number of bin states. Transition probabilities then describe the transition between the binned states.

Over any given interval k, the DER coordinator receives a number of charging and discharging requests. The coordinator then determines the proportion of charging and discharging requests to be accepted denoted as $\beta_{\rm c}[k]$ and $\beta_{\rm d}[k]$ respectively. This effect is achieved in the macromodel by transitioning the DERs from standby to corresponding charging and discharging states. Moreover, a timer tracks the total number of expiring packets $\beta_{\rm c}^-[k]$ and $\beta_{\rm d}^-[k]$ for past charging and discharging requests respectively to ensure transition back to standby states. The resulting dynamic is nonlinear and can be written as

$$q[k+1] = f(\beta[k], \beta^{-}[k], q[k]), \quad y[k] = h_{\text{dem}}(q[k])$$
 (4)

where, $\beta[k] = (\beta_c[k], \beta_d[k])$, $\beta^-[k] = (\beta_c^-[k], \beta_d^-[k])$, f(.) is a nonlinear map that describes aggregate PEM behavior. Finally, y[k] is the total power demand of the fleet and $h_{\text{dem}}(.)$ maps q[k] to output power. For full details of the macromodel, the reader is referred to [11].

It is worth mentioning here that the macromodel is a synchronous model of aggregate behavior of the fleet whereas requests are received asynchronously in real-time. Therefore, the macromodel and other synchronous models need to be validated in a real-time asynchronous simulation environment before deployment which is enabled by the platform developed and tested herein.

C. Baseline Power Consumption of a DER Fleet

The baseline power consumption of a fleet of DERs of the same load type is the minimum power signal (P_{nom}) for which the quality of service is guaranteed by PEM [10]. QoS is defined in terms of the average SoC of the fleet. For the case of electric water heaters, SoC is the average tank temperature of the fleet, scaled and shifted with respect to the dead-band and set-point. If the average temperature is close to the set-point, then we say that QoS is satisfied. Macromodel is used to obtain the baseline or nominal power consumption of the fleet and is obtained by solving the following optimization problem [12],

$$P_{\text{nom}}^* := \min_{\beta_{\text{c}}, \beta_{\text{d}} \in [0, 1]} \sum_{i=1}^{n} h_{\text{dem}}(q^*)$$
 subject to (5a)

$$q^* = f(\beta[k], \beta^-[k], q^*),$$
 (5b)

$$\left(q^*\right)^\top \chi \ge x_{\text{set}},\tag{5c}$$

where, x_{set} is the setpoint and χ consists of SoC associated with the binned state space corresponding to q[k]. The solution of the above optimization problem provides the total proportion of charging and/or discharging requests to be accepted $(\beta_c[k], \beta_d[k])$ that directly corresponds to a steady state distribution q^* in the macromodel and hence the baseline power consumption is obtained from $P_{\text{nom}} = h_{\text{dem}}(q^*)$.

The P_{nom} for DERs such as EWHs is driven by the average end-use consumption that is expected to vary depending upon the time of day. However, it is reasonable to assume that

average end-use consumption is relatively constant during a specific time period [31], [32]. For example, average end-use consumption and the corresponding P_{nom} during the afternoon period for EWHs, between 1pm and 5pm, can be assumed to be (relatively) constant. Thus, the results presented in this paper focus on a particular time period comprising of just 2 to 4 hours. However, a full-day simulation could be set up with time-varying end-use. Furthermore, end-use profiles are randomized and unique for individual EWHs in the real-time platform as detailed in [9] but with identical statistics within the time period of interest. The next section presents details regarding implementation of PEM enabled DERs on the grid-and-DER co-simulation platform and uses the macromodel to estimate SoC of the fleet.

IV. IMPLEMENTATION OF REAL-TIME DER COORDINATION

The objective of this section is to demonstrate that the cyberenabled, real-time simulator provides a realistic environment for testing advanced algorithms, such as the PEM approach described in the previous section. These tests increase confidence of DER coordination schemes before a real-world deployment.

A. Implementation of PEM on Simulator

The emulated DERs under PEM are executed as a thread in the processor's memory. Each DER then samples from the probability of request curve to determine if a request to consume a packet (charging or discharging but not both) is to be made to the coordinator. In case a request is made, the DER sends that request and then waits for the response. Requests are sent over the Internet using HTTP in Python and the coordinator responds to these requests using methods defined in requests module.

The coordinator, after receiving the request, either accepts or rejects it depending upon the total power consumption of the fleet and the power reference so that the tracking error is zero. That is, for the tracking error $P_{\text{error}}[k] := P_{\text{ref}}[k] - P_{\text{dem}}[k]$ a charge packet request with rated power P_1 is accepted if $P_{\text{error}}[k] + P_1 \le 0$ and rejected otherwise. On the other hand, a discharge packet with rated power P_{-1} is accepted only if $P_{\text{error}}[k] < 0$ and $P_{\text{error}}[k] + P_{-1} \ge 0$. It should be emphasized here that the coordinator is blind to the requesting DER's identity while making a decision to ensure privacy.

Each DER is equipped with a local clock to keep track of time elapsed since it started consuming a packet. As a result, asynchronism is ensured that is absent in standard software based simulations. In addition to the DER's local timer, the coordinator also associates a timer with each accepted packet request. This is achieved using Python's threading module. Timers allow the coordinator to keep track of the expiring packets and enables tighter tracking performance as shown in Section VI. It should also be noted here that the coordinator's ability to influence its decisions without explicit measurements such as DER's power is made possible only due to the PEM's

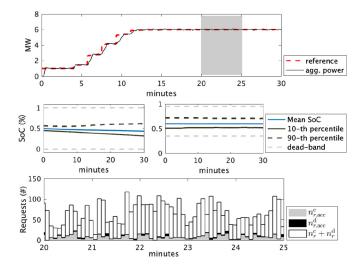


Fig. 3. This figure shows the tracking performance as well as the SoC of a fleet of heterogeneous DERs. The fleet consists of 4, 900 EWHs and 1, 150 ESSs emulated in real-time using the platform developed in this work. Two plots in the middle show the mean, 10-percentile and 90-th percentiles of the SoC of EWHs and ESS. Finally, the bottom plot shows the total number of accepted charging requests ($n_{r,acc}^c$) in black, total number of accepted discharging requests ($n_{r,acc}^d$) in grey color and aggregate charging and discharging requests made $n_r^c + n_r^d$. Since, requests are made asynchronously, therefore, total charging and discharging requests are added within each 5 second intervals for illustration in this plot.

unique request-response mechanism that acts as a natural feed-back. Furthermore, this feature is an advancement over the author's earlier implementation of the coordinator in [30].

1) Illustration With Real-Time Emulated DERs: Consider a heterogeneous fleet of diverse DERs consisting of 4,900 EWHs and 1,150 ESSs emulated in real-time using the developed simulator. The EWH parameters are Θ_{EWH} = $\{\mathcal{P}_{\zeta}^{\mathrm{EWH}}, \mathcal{L}, 52, [48.9, 55.1]\}, \ \mathcal{P}_{\zeta}^{\mathrm{EWH}} \sim \mathcal{N}(4.5, 0.25), \ \mathcal{L} \sim \mathcal{N}(200, 40)$ where $\mathcal{N}(\mu, \sigma)$ represents normal distribution with mean μ and standard deviation σ . Similarly, $\Theta_{\rm ESS} = \{\mathcal{P}_{\zeta}^{\rm ESS}, \mathcal{C}, 100, 75, [55, 95]\}, \ \mathcal{P}_{\zeta}^{\rm ESS} \sim \mathcal{N}(5, 0.5), \ \mathcal{C} \sim \mathcal{N}(13.5, 1).$ Grid is neglected in this example and P_{ζ} , η_{ζ} for charge and discharge modes are the same in ESSs. The fleet is coordinated by the aforementioned server that accepts or rejects requests to track a reference signal that varies from 1 MW to 6 MW in steps of 1 MW as shown in Fig. 3. Furthermore, the total number of accepted charging $(n_{r,acc}^c)$ and discharging $(n_{r,acc}^d)$ requests are also shown as the shaded region in a bar plot at the bottom of Fig. 3. Also included in the bar plot is the total number of charging and discharging requests $(n_r^c + n_r^d)$ made by the EWHs and ESSs respectively.

The average SoC of EWHs and ESS is plotted in Fig. 3 along with the 10-th and 90-th percentiles of the population during tracking. This shows that the SoC of ESS remains close to the set-point on average. Whereas EWH fleet is discharging as an aggregate which is due to the tracking reference being lower than the nominal obtained in the previous section. The average SoC of the fleet is an aggregate quantity which can be obtained using the macromodel as discussed next.

B. State of Charge of Aggregate

DERs are limited by the physical device constraints and can only be charged or discharged for a limited time. SoC of a fleet of DERs quantifies the flexibility available to the coordinator and is useful for making informed decisions. Let $x_i[k]$ be the SoC of the *i*-th DER, then the SoC of the fleet denoted as z[k] is defined as the average value of $x_i[k]$ normalized with respect to higher (\bar{x}) and lower (x) limits as follows,

$$z[k] := \frac{\sum_{i=1}^{N} \frac{x_i[k]}{N} - \underline{x}}{\overline{x} - \underline{x}}.$$
 (6)

Remark 2: Clearly, SoC (z[k]) in (6) requires individual state $x_i[k]$ to be measured. However, PEM does not measure each state, therefore, an estimate of z[k] is required. For this purpose, the extended Kalman filter (EKF) based on the macromodel has been proposed in [10] and can be used to estimate the SoC of the fleet. In this paper, it is assumed that groups of DERs of the same load type are aggregated and a SoC can be associated with each of those groups.

C. State of Charge Limits

The platform can now generate data available at the coordinator as well as that at DERs. One such use of data could be online data-driven characterization of fleet's SoC limits and will be considered in the future work. In this manuscript, however, we consider a simple model-based approach to determine limits of SoC and is summarized next.

In PEM, SoC of the fleet achieves it maximum value after authorizing all charging requests and rejecting all discharging requests for a long time. In terms of the macromodel, this is achieved by setting $\beta_c[k] = 1$ and $\beta_d[k] = 0$. By fixing the control signal $\beta[k] = (1,0)$ and the timer dynamics, (4) is reduced to an aperiodic, irreducible Markov chain which is guaranteed to have a stationary distribution. The maximum SoC limit, \bar{z} , is then the average SoC of the fleet corresponding to that distribution. Similarly, the minimum SoC limit is obtained by rejecting all charging requests ($\beta_c[k] = 0$) and accepting all discharging requests ($\beta_d[k] = 1$). The resulting stationary distribution of (4) gives the minimum limit \bar{z} . Furthermore, model-based SoC limits can be validated against the simple approach that uses the platform. The state estimation procedure using EKF is presented next.

D. State Estimation Using EKF

The Extended Kalman filter is used for state estimation (SE) that uses as measurements, i) the total requests, ii) aggregated power demand of the fleet and iii) opt-out rates, which are available to the coordinator. DERs that opt-out inform the coordinator using the same mechanism that is used for requests. A MySQL database table is also created to keep track of the power consumption of DERs. Each DER updates its current state in the table whenever the state changes. Aggregate power demand $P_{\rm dem}$ for EKF is then obtained from the database table by summing over all the entries corresponding to states of DERs. It should be mentioned here that individual power consumption of DERs is not used, only $P_{\rm dem}$. Furthermore, delays in measurement of $P_{\rm dem}$ can degrade

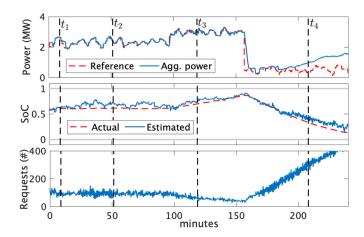


Fig. 4. Online estimation of the SoC (center) of DER aggregation consisting of 2,000 EWHs from measurements of total power consumption (top), number of request (bottom) and opt-outs only. DERs are shown to charge and discharge while tracking a scaled and shifted AGC signal.

tracking performance which can be improved by implementing timers at the coordinator as discussed in Section IV-A. Finally, the observability of PEM macromodel has previously been discussed in [10] where an EKF is designed. Standard EKF formulation is used herein as outlined in [10] and is illustrated in the next case study.

In this case study, the ability of the platform to support state estimation in real-time is illustrated for the simple case of 2,000 homogeneous EWHs with parameters $\Theta_{\rm EWH} = \{4.5,275,52,[48.9,55.1]\}$ under PEM. The baseline power consumption is obtained from the optimization problem (4) that results in $\beta_c^*[k] = 0.26$ and $P_{\rm nom}^* = 2.35$ MW. Note that since EWHs cannot discharge power into the grid as batteries, then the control signal reduces to the scalar $\beta[k] = \beta_c[k]$. The maximum and minimum energy limits are $53.4^{\circ}C$ and $49.3^{\circ}C$ respectively. Figure 4 shows EWHs tracking a scaled and shifted AGC signal.

Let t represent the simulation time, then from t = 0 min to t = 90 mins, the fleet's SoC is around nominal while tracking the AGC signal. This is because the AGC signal has been shifted to P_{nom}^* . Furthermore, the feedback from requests provide additional information about the SoC of the fleet. In this case, the requests rate on average is 100 requests per timestep which corresponds to a nominal request rate. It should be noted here that request rates depend upon the probability of request curve in PEM which is known to the coordinator since it is a design choice [9], [12].

The fleet is charging at an aggregate level from t = 90 mins to t = 160 mins since the reference is higher than P_{nom}^* . Request rate also decreases due to an increase in SoC. Finally, the fleet is discharging when the reference is below P_{nom}^* from t = 160 mins till the end of simulation. This is also indicated by an increase in the number of requests and the corresponding decrease in SoC. As the fleet discharges, number of DERs that opt-out of PEM increases. Recall that in opt-out, DERs do not make requests and simply consume power until SoC is sufficiently recovered. Therefore, P_{dem} increases as shown in the first subplot of Fig. 4.

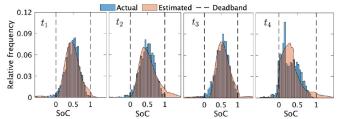


Fig. 5. Normalized temperature distribution of 2,000 EWHs at different times is shown in this plot. Estimated distribution from EKF in real-time (orange) matches closely with the actual distribution of EWHs (blue).

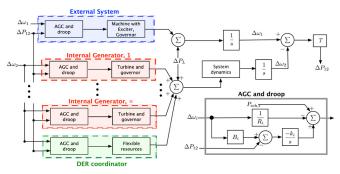


Fig. 6. Control schematic for the VELCO System, including internal and external generation and DER fleet. This dynamic grid model is simulated with OPAL-RT's ePHASORSIM.

In addition to the average SoC of the fleet, EKF also predicts the temperature distribution of EWHs as shown in Fig. 5. The temperature (SoC) distribution is shown corresponding to four different time instants from Fig. 4 marked as t_1 , t_2 , t_3 and t_4 . As illustrated in Fig. 4, EKF executed online accurately estimates the SoC in real-time from aggregate measurements only. Finally, this setup can also be used to tune the noise parameters of the EKF before real-world deployment.

V. CASE STUDY: PROVIDING GRID SERVICES WITH DERS

In this section, the real-time simulator is used to demonstrate that PEM resources can be used to manage regional mismatches between supply and demand in the grid under uncertain solar generation. Other grid services such as energy arbitrage, peak-load reduction etc. can also be tested but not shown in this paper due to space limitations.

A. Dispatching Flexible Resources in the Grid

Grid operators may pay high penalties for rescheduling generators or importing power through tie-lines to balance mismatches between supply and demand [33]. These power mismatches can be balanced by controlling flexible resources. Primary frequency regulation (speed-droop) on each generator and flexible resources can stabilize the power system with a steady state frequency deviation from the desired system frequency depending on the droop characteristic and frequency sensitivity [34]. Furthermore, automatic generation control (AGC) is included to drive the steady state frequency deviation to zero. Figure 6 shows the generalized control diagram for the system considered in this paper, which is

TABLE II
GRID AND DER PARAMETERS

Vermont Grid Parameters							
No. of buses/branches	161/223	Total load	609 MW				
Total renewables	305 MW	Cap. of external gen.	240 MW				
Cap. of local gen. 1	130 MW	Cap. of local gen. 2	35 MW				
Cap. of bulk battery	45 MWh	Nominal DER load	4.68 MW				

TABLE III
DROOP AND AGC PARAMETERS

Parameter	External	Local 1	Local 2	DER
R	33	33	33	20
B	1	1	1	1
k	1	1	1	5.1
$P_{\text{gen}}^{\text{sch}}\left(\text{MW}\right)$	218	86.2	26.6	4.68

an adapted version of the tie line control from [35]. A linear combination of frequency errors and change in imported power through tie-lines from their scheduled contract basis is used as an error signal called the area control error (ACE). AGC acts as a secondary control using an integral controller that sends out control signals to generators and the DER coordinator to reduce ACE to zero in steady state. For the purposes of this work, only two areas are used for simplicity, while being effective enough to demonstrate the flow of power between different areas. The first area is referred to as the "internal" area, representing the state of Vermont, whereas, the "external" area is used to represent the import and export of power from and to the Vermont transmission system. Furthermore, AGC gains of the DER coordinator are designed so that the contribution of DERs in frequency regulation decreases with the increase in the fleet's SoC. Next, the details of the grid used in case studies is described.

1) VELCO Test System: Vermont electric power company (VELCO) is Vermont's transmission system operator and provided the transmission grid data used to develop a realistic power system on which validation is performed. The ePHA-SORSIM model used in the simulator is then developed that consists of 161 buses, 135 branches, 89 transformers and 22 generation units that are lumped into two local generators labeled 1 and 2 respectively in Table II. Furthermore, power is imported and exported through tie-line interconnects and power imports are modeled as a lumped constant-power generator called external generator. The ePHASORSIM model is initialized based on VELCO data and AC load flow with reactive power limits on generators and a single slack bus. The generators are modeled in ePHASORSIM using the 6th order synchronous machine model. The exciter is modeled using the IEEE type ACA4 with the time constant 0.01 and the overall gain 200. The droop parameters (R), bias factors (B) and the integral gain of the generators as well as the DER fleet have been tuned according to their respective capacities. Finally, the PEM fleet has been modeled as a lumped constant power load. These parameters are given in Table III and an illustrative example using this model is presented next.

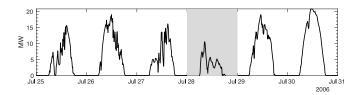


Fig. 7. Variation in solar generation over several days [36]. Case studies in this paper focuses on July 28, highlighted in grey color.

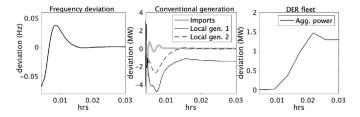


Fig. 8. The initial response of the system at the start of the simulation is shown in this figure. System frequency deviation returns to zero as the AGC dynamically adjusts local generator's and DER fleet's power output to match increasing solar generation.

B. Illustrative Example: Accounting for Uncertainty in Solar

Consider the scenario in which Vermont has time-varying distributed solar generation and flexible load consisting of 4,000 homogeneous EWHs. The base load of the fleet is 4.68 MW according to the nominal response of PEM described earlier and the corresponding SoC limits are given in Table II. Furthermore, Vermont has access to a bulk battery that can also be used. The objective of this bulk battery is to show that the DER coordinator can effectively co-optimize other resources along with flexible resources. Simulated solar power data for Vermont from [36] is used as shown in Figure. 7 for several days in July, 2006. The results presented in this paper focuses on a single day, July 28.

The initial frequency deviation and the deviation in generator's power output as a result of increased solar generation is plotted in Fig. 8 while the system performance over the next hour and a half is plotted in Fig. 9. AGC ensures that the power imported through the tie-line remains close to the scheduled value by dynamically modifying the setpoints of local generators, the bulk battery as well as the flexible EWHs under PEM coordinator as shown in Fig. 9. As a result, the EWH fleet starts to charge and its SoC increases. When the SoC gets closer to the upper limit, the AGC reduces the power setpoint of the PEM coordinator. It should also be noted that by using flexible resources, the more expensive imports remain fixed at their scheduled values whereas the less expensive EWHs, bulk battery and local generators compensate for the excess solar generation.

VI. INCORPORATING DER-RELATED PRACTICAL LIMITATIONS IN THE REAL-TIME SIMULATOR

This section shows that the real-time platform allows the coordinator to identify potential practical limitations in real-world deployment. Two cases are considered; (i) effect of measurement delays and (ii) delays introduced due to latency in the communication channel. The results presented here uses

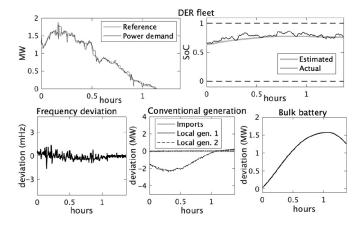


Fig. 9. Top row shows the DER fleet tracking a reference signal generated by the AGC based on its SoC. The AGC is designed to account for the SoC of the fleet and as the SoC increases, the tracking power reference is lowered. Power reference and consumption is plotted in terms of deviation from the nominal consumption. Bottom row shows the frequency deviation of the system in mHz as well as the variation in power from conventional generation and total imported power. The imported power remains fixed at its scheduled value whereas the bulk battery and the less expensive local generation compensates for the excess solar generation.

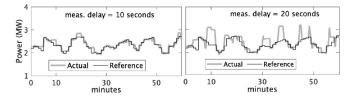


Fig. 10. This plots shows the effect of measurement delays in $P_{\rm dem}$ on tracking performance of DERs. The right plot is for the case when the average measurement delay is 20 seconds resulting in RMS tracking error of 60kW. Similarly, the left plot shows the case in which $P_{\rm dem}$ measurements are delayed by 30 seconds on average. RMS tracking error in this case is 160.6kW.

PEM for demonstration purposes, however, other schemes can also be implemented.

A. Effect of Measurement Delays

Delays in the measurement of aggregate power $P_{\rm dem}$ are studied first which can be delayed due to practical limitations. The coordinator then uses the most recent measurement of $P_{\rm dem}$ to decide the number of requests to be accepted. It should be emphasized here that the coordinator is making these decisions in real-time and in a sequential manner. Recall that a DER once allowed to consume a packet will either charge or discharge for the entire duration of the packet length. Delays in the power measurement directly influences the performance of the DER coordinator since authorizing incorrect number of packets can not be cancelled under PEM. This is because the coordinator is unaware of the identity of the DER associated with each request that further ensures privacy of the participating DER.

To study the effect of delays, it is assumed that the probability of a $P_{\rm dem}$ measurement being delayed is 10%. Delay time itself is distributed according to a normal distribution with mean $\{20, 30, 60\}$ seconds and a standard deviation of 2 seconds. Fig. 10 shows that a 20 second average delay introduces

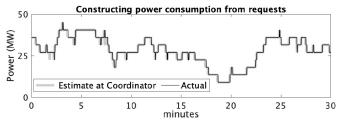


Fig. 11. This figure shows that the DER coordinator can use the requests and a timer mechanism to construct power consumption estimate and is demonstrated here for 20 EWHs.

an RMS tracking error of about 60kW which translates to 2.5% with respect to the baseline power consumption. The maximum RMS error of 15% is observed in the extreme case of 60 seconds. In reality, delays are usually less than 20 seconds and Fig. 10 shows that PEM resources can deliver excellent tracking in real-time even in the presence of delays. However, in PEM the coordinator can also make use of the packet request mechanism to enhance the tracking performance as discussed next.

1) Constructing Real-Time Power Demand of DER Fleet: The inherent packet based mechanics in PEM allows the coordinator to approximate the aggregate power consumption of the fleet in real-time. Consider first the charging requests and recall that accepting a request corresponds to an increase in the aggregate power P_{dem} by P_1 , whereas, packet expiration corresponds to a decrease in power by P_1 . Therefore, at any time instant, P_{dem} of the fleet increases according to the total number of DERs authorized by the coordinator to charge and the total number of DERs opting out. Both these quantities are known to the coordinator as mentioned in Section III. In a homogeneous population of DERs, P_1 is the same for each packet. For heterogeneous DERs, an additional rated power field in the request can be added without extra communication overheard. In order to keep tracking of expiring packets, the coordinator associates each packet with a timer equal to the packet length δ_1 as mentioned in Section IV-A. As the timer expires, the coordinator models that the associated DER has transitioned to standby and adjusts P_{dem} accordingly. This packet based timer mechanism allows the coordinator to construct P_{dem} when measurements are delayed.

The plot in Fig. 11 shows that the estimated power matches that of the true aggregate power of the fleet ($P_{\rm dem}$). Only 20 EWHs are considered to demonstrate the accuracy of these measurements. These concepts are then easily extended to the case of discharging packets. DER coordinator operating under PEM is therefore equipped with a built-in feedback mechanism in terms of requests that allows tight tracking performance and accurate SoC estimation observed in the earlier section.

B. Effect of Input Delays

Input delays are incurred due to the time between the DER coordinator sending a charge/discharge authorization and the corresponding DER switching its state. To understand the source of such delays, consider the sequence of events in PEM's request mechanism, conceptually represented in

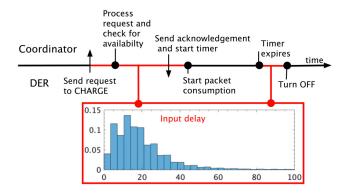


Fig. 12. (Top) Sequence of events that occur between the aggregator and the DERs in PEM's request mechanism, (bottom) measured input delay between request acceptance and packet consumption in PEM.

Fig. 12. The DER sends a request to consume a packet to the coordinator which then responds with an acknowledgement and initiates a timer. Whereas, the DER waiting for the response, starts to consume the packet only after the acknowledgement has arrived. Furthermore, when the timer at the coordinator expires, it assumes that P_{dem} should be reduced. However, it is possible that the DER's own dynamics might extend this packet consumption time and lead to an additional delay. For the developed cyber-enabled real-time simulator, the combination of both these delays is distributed as shown in Fig. 12 with a mean of 64ms for 2,000 EWHs. RMS error between the actual power consumption and the estimated power consumption is 13.06kW. In real-world deployment of [30], these delays are observed to be 200ms on average. To study the effect of increased delays on the DER coordinator's estimates of power consumption, input delays are artificially injected. This results in RMS error of 35.4kW which is only about 1.7% of the nominal power consumption and resulted in good tracking performance.

VII. CONCLUSION

This manuscript presents a real-time grid-and-DER co-simulation platform that enables and characterizes performance evaluation of large-scale DER coordination schemes with asynchronous communications in a realistic setting. Importantly, to support validation efforts and lower barriers to real-world DER program deployments, this platform captures salient cyber and physical limitations, including communication systems, device behaviors, and grid challenges. Moreover, the co-simulator is agnostic to the type of coordination scheme and allows real-time implementation of model-based state-estimation and control algorithms. Illustrative results from the co-simulation platform are presented for the DER demand dispatch scheme called PEM. PEM has a unique request-response mechanism that coordinates DERs via bi-directional DER-to-coordinator communications, which makes it well suited for exploring the capability of the real-time grid-and-DER co-simulation platform.

Future work seeks to implement and compare different centralized and distributed coordination schemes in terms of performance and QoS guarantees. Furthermore, since different types of DERs operating under the same coordinator are suitable for several applications such as peak-load reduction, frequency regulation etc., therefore, we are interested in application-specific grouping and validation of diverse DER fleets.

ACKNOWLEDGMENT

The authors would like to thank Andrew Klem, Prof. Jeff Frolik from University of Vermont and Prof. Sumit Paudyal from Florida International University for valuable discussions on PEM based real-time grid simulations and experiments. Furthermore, they are grateful to Bernadette Fernandes and Chris Root of VELCO for insightful discussions and providing transmission and bus data of the Vermont system to create a realistic case study. P. D. H. Hines and M. R. Almassalkhi are co-founders of startup Packetized Energy, which commercialized a version of Packetized Energy Management.

REFERENCES

- M. Amini and M. Almassalkhi, "Trading off robustness and performance in receding horizon control with uncertain energy resources," in *Proc. Power Syst. Comput. Conf.*, Jun. 2018, pp. 1–7.
- [2] F. C. Schweppe, R. D. Tabors, J. L. Kirtley, H. R. Outhred, F. H. Pickel, and A. J. Cox, "Homeostatic utility control," *IEEE Trans. Power App. Syst.*, vol. PAS-99, no. 3, pp. 1151–1163, May 1980.
- [3] M. G. Morgan and S. N. Talukdar, "Electric power load management: Some technical, economic, regulatory and social issues," *Proc. IEEE*, vol. 67, no. 2, pp. 241–312, Feb. 1979.
- [4] R. Malhamé, "A jump-driven Markovian electric load model," Adv. Appl. Probab., vol. 22, no. 3, pp. 564–586, 1990.
- [5] D. S. Callaway, "Tapping the energy storage potential in electric loads to deliver load following and regulation, with application to wind energy," *Energy Convers. Manag.*, vol. 50, no. 5, pp. 1389–1400, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/ pii/S0196890408004780
- [6] S. P. Meyn, P. Barooah, A. Bušić, Y. Chen, and J. Ehren, "Ancillary service to the grid using intelligent deferrable loads," *IEEE Trans. Autom. Control*, vol. 60, no. 11, pp. 2847–2862, Nov. 2015.
- [7] J. L. Mathieu, M. Kamgarpour, J. Lygeros, and D. S. Callaway, "Energy arbitrage with thermostatically controlled loads," in *Proc. Eur. Conf. Circuit Theory Design*, Jul. 2013, pp. 2519–2526.
- [8] L. C. Totu and R. Wisniewski, "Demand response of thermostatic loads by optimized switching-fraction broadcast," *IFAC Proc. Vol.*, vol. 47, no. 3, pp. 9956–9961, 2014. [Online]. Available: http://www.science direct.com/science/article/pii/S1474667016431927
- [9] M. Almassalkhi, L. D. Espinosa, P. D. Hines, J. Frolik, S. Paudyal, and M. Amini, "Asynchronous coordination of distributed energy resources with packetized energy management," in *Energy Markets and Responsive Grids*, vol. 162. New York, NY, USA: Springer, 2018, pp. 333–361.
- [10] L. A. D. Espinosa, M. Almassalkhi, P. Hines, and J. Frolik, "System properties of packetized energy management for aggregated diverse resources," in *Proc. Power Syst. Comput Conf.*, Jun. 2018, pp. 1–6.
- [11] L. A. D. Espinosa and M. Almassalkhi, "A packetized energy management macromodel with quality of service guarantees for demand-side resources," *IEEE Trans. Power Syst.*, vol. 35, no. 5, pp. 3660–3670, Sep. 2020.
- [12] L. A. D. Espinosa, A. Khurram, and M. Almassalkhi, "Reference-tracking control policies for packetized coordination of heterogeneous DER populations," *IEEE Trans. Control Syst. Technol.*, vol. 29, no. 6, pp. 2427–2443, Nov. 2021.
- [13] D. Serpanos, "The cyber-physical systems revolution," Computer, vol. 51, no. 3, pp. 70–73, Mar. 2018.
- [14] S. Xin, Q. Guo, H. Sun, B. Zhang, J. Wang, and C. Chen, "Cyber-physical modeling and cyber-contingency assessment of hierarchical control systems," *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2375–2385, Sep. 2015.

- [15] M. Ayar, R. D. Trevizan, S. Obuz, A. S. Bretas, H. A. Latchman, and N. G. Bretas, "Cyber-physical robust control framework for enhancing transient stability of smart grids," *IET Cyber Phys. Syst. Theory Appl.*, vol. 2, no. 4, pp. 198–206, 2017.
- [16] M. P. Korukonda, S. R. Mishra, K. Rajawat, and L. Behera, "Hybrid adaptive framework for coordinated control of distributed generators in cyber-physical energy systems," *IET Cyber Phys. Syst. Theory Appl.*, vol. 3, no. 1, pp. 54–62, 2018.
- [17] A. Pratt, M. Baggu, F. Ding, S. Veda, I. Mendoza, and E. Lightner, "A test bed to evaluate advanced distribution management systems for modern power systems," in *Proc. IEEE EUROCON 18th Int. Conf. Smart Technol.*, 2019, pp. 1–6.
- [18] A. Hahn, A. Ashok, S. Sridhar, and M. Govindarasu, "Cyber-physical security testbeds: Architecture, application, and evaluation for smart grid," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 847–855, Jun. 2013.
- [19] D. Nicol, C. Davis, and T. Overbye, "A virtual power system testbed for cyber-security decision support," in *Proc. INFORMS Simulat. Soc. Workshop Simulat. Interface Model. Anal.*, 2009, pp. 62–66.
- [20] K. Kok et al., "Dynamic pricing by scalable energy management systems—Field experiences and simulation results using PowerMatcher," in Proc. IEEE Power Energy Soc. Gen. Meeting, 2012, pp. 1–8.
- [21] J. Johnson, R. Ablinger, R. Bründlinger, B. Fox, and J. Flicker, "Design and evaluation of SunSpec-compliant smart grid controller with an automated hardware-in-the-loop testbed," *Technol. Econ. Smart Grids Sustain. Energy*, vol. 2, no. 1, pp. 1–10, 2017.
- [22] Z. Taylor, H. Akhavan-Hejazi, and H. Mohsenian-Rad, "Optimal operation of grid-tied energy storage systems considering detailed device-level battery models," *IEEE Trans. Ind. Informat.*, vol. 16, no. 6, pp. 3928–3941, Jun. 2020.
- [23] U.S. DoE Office of Electricity (OE). "Transactive energy simulation platform, version 00." Mar. 2017. [Online]. Available: https://www.osti.gov//servlets/purl/1348185
- [24] R. Melton. "Pacific northwest smart grid demonstration project technology performance report volume 1: Technology performance." Jun. 2015. [Online]. Available: https://www.osti.gov/biblio/1367568
- [25] M. D. Ilić, R. Jaddivada, and X. Miao, "Scalable electric power system simulator," in *Proc. IEEE PES Innov. Smart Grid Technol. Conf. Europe* (ISGT-Europe), 2018, pp. 1–6.
- [26] D. M. Nicol, C. M. Davis, and T. Overbye, "A testbed for power system security evaluation," *Int. J. Inf. Comput. Security*, vol. 3, no. 2, pp. 114–131, 2009.
- [27] H. Neema, W. Emfinger, and A. Dubey, "A reusable and extensible Web-based co-simulation platform for transactive energy systems," in *Proc. 3rd Int. Trans. Energy Syst.*, vol. 12. Portland, OR, USA, 2016, pp. 1–8.
- [28] G. S. Ledva, E. Vrettos, S. Mastellone, G. Andersson, and J. L. Mathieu, "Managing communication delays and model error in demand response for frequency regulation," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1299–1308, Mar. 2018.
- [29] M. Amini, A. Khurram, A. Klem, M. Almassalkhi, and P. D. Hines, "A model-predictive control method for coordinating virtual power plants and Packetized resources, with hardware-in-the-loop validation," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, 2019, pp. 1–5.
- [30] K. Desrochers et al., "Real-world, full-scale validation of power balancing services from packetized virtual batteries," in Proc. IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf. (ISGT), Feb. 2019, pp. 1–5.
- [31] A. Khurram, L. A. D. Espinosa, M. Almassalkhi, and R. Malhame, "Identification of hot water end-use process of electric water heaters from energy measurements," in *Proc. Power Syst. Comput. Conf.*, Jun. 2020, pp. 1–8.
- [32] A. Khurram, L. A. D. Espinosa, and M. R. Almassalkhi, "A methodology for quantifying flexibility in a fleet of diverse DERs," in *Proc. IEEE Madrid PowerTech*, 2021, pp. 1–6.
- [33] M. Arnold and G. Andersson, "Model predictive control of energy storage including uncertain forecasts," in *Proc. Power Syst. Comput. Conf.*, vol. 23, 2011, pp. 24–29.
- [34] M. Amini and M. Almassalkhi, "Investigating delays in frequency-dependent load control," in *Proc. IEEE Innov. Smart Grid Technol. Asia (ISGT-Asia)*, Nov. 2016, pp. 448–453.
- [35] P. Kundur, N. J. Balu, and M. G. Lauby, Power System Stability and Control, vol. 7. New York, NY, USA: McGraw-Hill, 1994.
- [36] National Renewable Energy Laboratory (NREL). "Solar power data for integration studies." [Online]. Available: https://www.nrel.gov/grid/solarpower-data.html (Accessed: Jan. 16, 2020).



Adil Khurram (Member, IEEE) received the B.S. degree in electrical engineering from the National University of Sciences and Technology, Islamabad, Pakistan, in 2013, the M.S. degree in electrical engineering from the American University of Sharjah, Sharjah, UAE, in 2016, and the Ph.D. degree in electrical engineering from the University of Vermont, Burlington, VT, USA. He is currently working as a Postdoctoral Researcher with the University of California at San Diego, San Diego, CA, USA, on the NSF funded project DERConnect. His research

interests include statistical electrical load modeling and control of distributed energy resources in power systems.



Mahraz Amini (Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical and computer engineering from the University of Tehran, Iran, in 2011 and 2014, respectively, and the Ph.D. degree in electrical engineering from the University of Vermont in 2019. He is currently working as a Lead Engineering Analyst with National Grid, Waltham, MA, USA. His research interests include power system analysis, stochastic optimization, and control in power systems. His recent activity has focused largely on integration of renewable generation and energy

storage systems into power systems.



Luis A. Duffaut Espinosa (Senior Member, IEEE) received the B.S. degree in physics from the Universidad Nacional de Ingeniería, Lima, Perú, in 2003, the M.S. degree in mathematics with mention in stochastic processes from the Pontificia Universidad Católica del Perú, Lima, Perú, in 2005, and the Ph.D. degree in electrical and computer engineering form Old Dominion University, Norfolk, VA, USA, in 2009. He is currently an Assistant Professor with the Department of Electrical and Biomedical Engineering, University of Vermont,

Burlington, VT, USA. His research interests span modeling, control, and estimation of nonlinear systems, stochastic processes, and algebraic combinatorics.



Paul D. H. Hines (Senior Member, IEEE) received the B.S. degree in electrical engineering from the University of Washington, Seattle, WA, USA, in 1997, the M.S. degree in electrical engineering from Seattle Pacific University, Seattle, in 2001, and the Ph.D. degree in engineering and public policy from Carnegie Mellon University, Pittsburgh, PA, USA, in 2007. He is currently the Vice President of Power Systems with EnergyHub. Previously, he was a Professor with the Department of Electrical and Biomedical Engineering, University of Vermont,

Burlington, VT, USA, and the Co-Founder of Packetized Energy.



Mads R. Almassalkhi (Senior Member, IEEE) received the B.S. degree in electrical engineering with a dual major in applied mathematics from the University of Cincinnati, Cincinnati, OH, USA, in 2008, and the M.S. degree in electrical engineering: systems and the Ph.D. degree from the University of Michigan at Ann Arbor, Ann Arbor, MI, USA, in 2010 and 2013, respectively. He is currently an Associate Professor with the Department of Electrical and Biomedical Engineering, University of Vermont, Burlington, VT, USA, and the Chief

Scientist with Pacific Northwest National Laboratory (by joint appointment). His research interests span control of DERs, grid optimization, and multi-energy systems. He is also an Associate Editor of IEEE TRANSACTIONS ON POWER SYSTEMS.