Enhancing the Spatio-Temporal Observability of Grid-Edge Resources in Distribution Grids

Shanny Lin , *Graduate Student Member, IEEE*, and Hao Zhu , *Senior Member, IEEE*

Abstract—Enhancing the spatio-temporal observability of distributed energy resources (DERs) is crucial for achieving secure and efficient operations in distribution grids. This paper puts forth a joint recovery framework for residential loads by leveraging the complimentary strengths of heterogeneous measurements in real time. The proposed framework integrates low-resolution smart meter data collected at every load node with fast-sampled feeder-level measurements from limited number of distribution phasor measurement units. To address the lack of data, we exploit two key characteristics for the loads and DERs, namely the sparse changes due to infrequent activities of appliances and electric vehicles (EVs) and the locational dependence of solar photovoltaic (PV) generation. Accordingly, meaningful regularization terms are introduced to cast a convex load recovery problem, which will be further simplified to reduce the computational complexity. The load recovery solutions can be utilized to identify the EV charging events at each load node and to infer the total behind-the-meter PV output. Numerical tests using real-world data have demonstrated the effectiveness of the proposed approaches in enhancing the visibility of grid-edge DERs.

Index Terms—Distributed energy resources, synchrophasor data, smart meter, distribution state estimation, subspace learning.

I. Introduction

DISTRIBUTION grids have witnessed an increasing penetration of distributed energy resources (DERs) at the grid-edge. Electrical vehicles (EVs) and renewable generation owned by residential customers and aggregators lead to growing dynamics in distribution systems. High-resolution dynamic profiles of residential loads have been advocated for validating EV charging commands [1] and photovoltaic (PV) inverter control settings [2]. In addition, the instantaneous total heating/cooling loads and solar PV generation within a feeder are useful for designing protection systems and demand response programs [3], [4]. Therefore, it is important to enhance the spatio-temporal observability of grid-edge DERs for efficient and secure grid operations.

Nonetheless, distribution systems traditionally lack in sensing infrastructure and real-time observability. To address

this issue, different types of distribution sensors have been advocated recently. At the feeder level, line flow meters and distribution phasor measurement units (D-PMUs) [5] can provide fast power measurements of high quality, but only at selected locations. Meanwhile, almost all residential homes are equipped with smart meters that can collect the electricity consumption data over intervals of 15 minutes to an hour [6]. Hence, high resolution D-PMU data can suffer from low spatial diversity, while the ubiquitous smart meter data lack in the temporal resolution. These two types of measurements provide complimentary strengths, and the question opens up on how to jointly utilize them both to enhance the visibility of grid-edge DERs.

The present work is highly related to the distribution state estimation (DSE) problem. Traditional DSE methods [7], [8] have focused on the static setting and identified the lack of measurements as the main challenge. To tackle this, more recent DSE methods have considered the coupling between two consecutive time instances [9] and the spatial correlation among voltage and power data [10]. All of these approaches are not directly applicable to the case of different measurement time resolutions. Motivated by the availability of fast D-PMU data, dynamic DSE methods have been developed by building on Kalman filtering [11], [12] or prediction-correction method [13], including a recent extension to asynchronous data in [14]. These recursive update based methods can well track slowly varying system changes but may fail to cope with fast start/end events that are common for residential appliances and DERs. In addition, pseudo measurements using load forecast have been popularly considered for DSE, such as the Bayesian estimation approach in [15]. Nonetheless, it is very difficult to predict appliance activities from historic data. See [16], [17] for comprehensive reviews of the DSE literature. Another line of related work is the load disaggregation for inferring behind-the-meter (BTM) PV output. However, these approaches typically require prior knowledge such as historical load data for a subset of customers or the statistical models for the total consumption; see, e.g., [3], [18] and references therein. Such information can be difficult to obtain while most of these approaches only utilize slow smart meter data. Thus, the problem still remains on how to effectively incorporate dynamic data streams of various resolutions to enhance the observability of grid-edge DERs.

The goal of this paper is to develop a joint spatio-temporal DER inference framework by using the respective strengths of D-PMU and smart meter data. Both types of data are first modeled as linear transformation of the unknown power

demands. This model builds upon linearizing the multiphase power flow equations and generalizes our earlier work [19] for lossless systems. To tackle the lack of measurements, two key underlying characteristics of the spatio-temporal load demand are exploited. First, the on-off activities of household appliances and EVs typically occur *infrequently*, leading to *jointly sparse changes* of both active and reactive powers. Second, renewable generation such as rooftop solar PV demonstrates a strong *spatial correlation* within a feeder, giving rise to a *low-rank* component in the active power demand. Note that the correlation of localized PV outputs here is different from earlier work [20], [21] on low-rankness of the load data. These approaches exploit the relation among various household appliances for disaggregating the total house load consumption. To promote the low-rank plus sparse change feature, group L1-norm [22] and nuclear norm [23]–[25] are introduced as regularization, leading to a convex formulation. The low-rank component is further simplified into a rank-one PV component to reduce the computational complexity associated with nuclear norm. Numerical tests using real-world load and DER data have been conducted to demonstrate the effectiveness of the joint inference formulations in identifying EV events and estimating the total BTM solar output in a feeder. We have observed that a few D-PMUs can significantly improve the DER visibility than only using smart meter data, while the presence of periodic appliance activities can affect the performance of DER inference.

The main contribution of this work is summarized here. First, we consider the inference problem for monitoring the dynamics of residential DERs that can incorporate distribution system measurements of different time resolutions and spatial availability in Section II. Second, the proposed spatio-temporal inference framework has exploited the underlying low-rank and sparse-change characteristics to improve the recovery performance of residential DERs in Section III. Last, the proposed methods have been successfully applied to crucial grid-edge monitoring tasks such as identifying the exact timing of EV events and estimating the BTM solar generation in Section IV. Conclusions and future directions are drawn in Section V.

Notation: Upper (lower) boldface symbols stand for matrices (vectors); $(\cdot)^{\mathsf{T}}$ stands for matrix transposition; $(\cdot)^*$ complex conjugate; $\|\cdot\|_*$ denotes the matrix nuclear norm; $\|\cdot\|_g$ the group L1-norm; $\|\cdot\|_2$ the L2-norm; $\otimes$ denotes the Kronecker product; $\mathbf{I}$ ($\mathbf{1}$) stands for the identity matrix (all-one vector) of appropriate size; $\mathrm{diag}(\cdot)$ represents the diagonal matrix; and symbols using $\check{}$ indicate vectors of all non-reference bus quantities.

II. SYSTEM MODEL

Consider a multiphase distribution feeder with residential loads at the feeder ends shown by Fig. 1. We are interested in the spatio-temporal complex load matrix $(\mathbf{P}+\mathbf{jQ}) \in \mathbb{C}^{N \times T}$, or the real-valued version $\mathbf{X} := [\mathbf{P}; \mathbf{Q}] \in \mathbb{R}^{2N \times T}$, where N denotes the number of single-phase loads and T the number of time slots. The temporal resolution represents the fastest sampling rate of all available measurements. For example, the
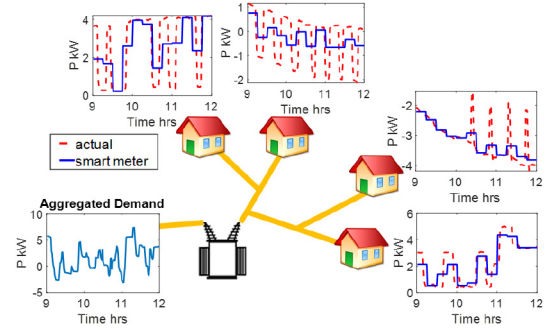

Fig. 1. Overview of a simple distribution feeder system with four residential loads and various types of power measurements.

sampling rate can be at the sub-second scale which corresponds to the time resolution of D-PMUs [5]. Without loss of generality (Wlog), this paper uses the fast sampling rate at every minute.

Two types of measurement data will be mainly utilized, namely the smart meter data for each load and the aggregated feeder-level measurements. Residential houses are equipped with smart meters to record the average electricity consumed over intervals of 15 minutes to an hour [6]. Here we assume both real and reactive powers are collected every 15 minutes. Using this down-sampling rate to obtain $T_s = T/15$ samples, one can form the smart meter data matrix of size $2N \times T_s$ as

$$\mathbf{\Gamma} = \begin{bmatrix} \mathbf{\Gamma}^P \\ \mathbf{\Gamma}^Q \end{bmatrix} = \begin{bmatrix} \mathbf{P} \\ \mathbf{Q} \end{bmatrix} \mathbf{A} + \mathbf{N}^{\mathbf{\Gamma}} \qquad (1)$$

where the $T \times T_s$ matrix $\mathbf{A} = (1/15)\mathbf{I} \otimes \mathbf{1}$ is the time averaging operator, while $\mathbf{N}^{\mathbf{\Gamma}}$ denotes the measurement noise matrix. We extend (1) to the case of asynchronized sampling, by using a node-specific averaging $\mathbf{A}_n$ according to its sampling intervals. As shown by the numerical tests later on, this generalization can be easily included by the rest of analysis as the measurement model is still linear.

At the feeder level, D-PMUs or other meters installed on power lines can measure quantities such as aggregated demand at a fast rate. To model the feeder-level data, we briefly introduce the multiphase ac power flow equations; see, e.g., [26] for a comprehensive overview. Let the substation or feeder head be the reference bus 0, and use N_b denote the number of other buses (PQ buses). Wlog, all the buses are three-phase connected, and for simplicity, only wye-connected single-phase loads are considered. Let vectors $\check{\mathbf{s}} \in \mathbb{C}^{3N_b}$ and $\check{\mathbf{v}} \in \mathbb{C}^{3N_b}$ collect the complex power injections and phase-to-ground voltages at every phase of the PQ buses. The power injection $\check{\mathbf{s}}$ is quadratically related to the voltage $\check{\mathbf{v}}$, as

$$\check{\mathbf{s}} = \mathrm{diag}(\check{\mathbf{v}})\check{\mathbf{i}}^* = \mathrm{diag}(\check{\mathbf{v}})(\mathbf{Y}_L[\mathbf{v}_0; \check{\mathbf{v}}])^* \qquad (2)$$

where $\mathbf{v}_0 \in \mathbb{C}^3$ is the given reference voltage vector at feeder head, while matrix $\mathbf{Y}_L \in \mathbb{C}^{3N_b \times 3(N_b+1)}$ is the submatrix of the full admittance (Ybus) matrix $\mathbf{Y} = [\mathbf{Y}_0; \mathbf{Y}_L] \in \mathbb{C}^{3(N_b+1) \times 3(N_b+1)}$ that consists of rows of all non-reference PQ buses. Similarly, the power injection to the reference bus $\mathbf{s}_0 \in \mathbb{C}^3$, or the total aggregated power demand, becomes

$$\mathbf{s}_0 = \mathrm{diag}(\mathbf{v}_0)\mathbf{Y}_0^*[\mathbf{v}_0; \check{\mathbf{v}}]^* \qquad (3)$$

where $\mathbf{Y}_0 \in \mathbb{C}^{3 \times 3(N_b+1)}$ is the submatrix of Ybus with rows corresponding to the reference bus. In general, $\mathbf{s}_0$ and the power/current flow elsewhere in the feeder are fully determined by the voltage profile $\check{\mathbf{v}}$.

To simplify the solution of $\check{\mathbf{v}}$, one can adopt linearized power flow approximations such as the fixed-point method (FPM) developed in [27], [28]. Consider the linearized version of (2) given by

$$\tilde{\mathbf{v}} = \mathbf{M}\check{\mathbf{x}} + \check{\mathbf{w}} \qquad (4)$$

where $\check{\mathbf{x}} := [\mathsf{Re}\{\mathbf{s}\}; \mathsf{Im}\{\mathbf{s}\}] \in \mathbb{R}^{6N_b}$ has the full nodal injection and $\check{\mathbf{w}} \in \mathbb{C}^{3N_b}$ is the complex voltage solution under zero-loading condition. Basically, the FPM-based approximation is the linear interpolation of $\check{\mathbf{w}}$ and another voltage solution (hence the term fixed-point). Substituting the linearized voltage model (4) into (3) yields

$$\tilde{\mathbf{s}}_0 = \mathsf{diag}(\mathbf{v}_0)\mathbf{Y}_{0L}^*\mathbf{M}^*\check{\mathbf{x}} \qquad (5)$$

where $\mathbf{Y}_{0L} \in \mathbb{C}^{3 \times 3N_b}$ is the submatrix of $\mathbf{Y}_0 = [\mathbf{Y}_{00}, \mathbf{Y}_{0L}]$ formed by the PQ-bus columns only. Note that the additional offset term is canceled in (5) so that $\tilde{\mathbf{s}}_0 = \mathbf{0}$ under the zero-loading condition. Wlog, let $\check{\mathbf{x}}$ include only the load nodes of non-zero injection, and thus it reduces to a $(2N \times 1)$ vector $\mathbf{x}$ which matches the dimension of matrix $\mathbf{X}$. By stacking (5) across T time slots, we obtain the measurement model for the aggregated demand as

$$\mathbf{Z} = \begin{bmatrix} \mathbf{Z}^P \\ \mathbf{Z}^Q \end{bmatrix} = \mathbf{HX} + \mathbf{N}^Z \in \mathbb{R}^{6 \times T} \qquad (6)$$

where $\mathbf{N}^Z$ is the measurement error matrix.

Remark 1 (Measurement types and accuracy): The FPM-based measurement model (6) can encompass a variety of feeder-level quantities in addition to aggregated power, such as the data on line power/current flow and nodal voltage phasor/magnitude as measured by D-PMUs; see more discussions in [16]. Intuitively, the larger the amount of available data, the better the performance a recovery algorithm would attain. For example, our numerical experiences (not included due to space limit) confirm that the proposed joint recovery framework using both D-PMU and smart meter data clearly outperforms the recovery using the latter only. As for the accuracy of the FPM-based linear approximation, it depends on the fixed power flow solution used to form $\mathbf{M}$, as shown in [27], [28]. Ideally, this fixed point should closely approach the actual power flow solution. Using the latest smart meter data, one can update matrix $\mathbf{M}$ of (5) for improved accuracy of linearization. As the aggregated demand model (6) would still be linear in $\mathbf{X}$, this time-varying $\mathbf{M}$ scenario can be easily considered by the rest of analysis as shown by our numerical tests later on.

Even though encompassing various measurements, the problem of recovering $\mathbf{X}$ is expected to be severely under-determined. For the measurement models in (1) and (6), the total number of equations equals $(2NT_s + 6T)$ which is much smaller than $(2NT)$, the number of unknowns. This motivates one to exploit the special characteristics of the load matrices for improving the recovery performance.
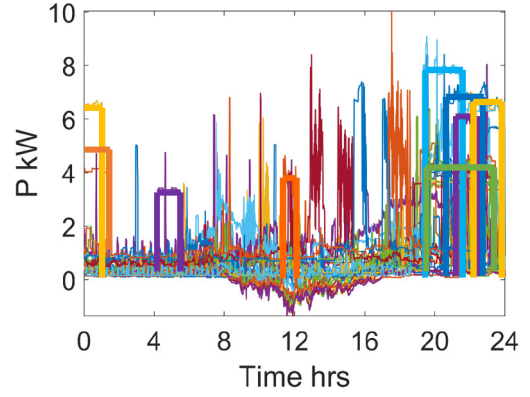

Fig. 2. Residential load profiles of active power from Pecan Street's Dataport [32] for 30 houses, 15 of which have rooftop PVs.

III. MONITORING DERs VIA LOAD RECOVERY

We first discuss the characteristics of residential load matrices that are amendable for formulating the load recovery problem. Typically, residential load profiles consist of the base loading and household appliance activities; see, e.g., [29]. Hence, each residential load can be approximated by the aggregation of its own rectangular waveforms that are independent from other locations. DERs such as EVs are operated as household appliances of large power ratings, and thus their load profiles also follow the rectangular waveform condition. For example, Fig. 2 plots the active power profiles of 30 actual residential homes, of which the large night-time rectangular waveforms correspond to EV charging activities. A majority of appliance activities are observed to occur *infrequently* throughout the day, and have a non-unity inductive power factor [30]. Hence, they can be represented as *sparse changes* in the household load profile, in both active and reactive power. Notice that this sparse-change characteristic was also exploited in [9] for static distribution system state estimation. Some periodic appliances such as HVACs may violate this condition, as discussed later in Remark 5.

To this end, let $\mathbf{S}^P, \mathbf{S}^Q \in \mathbb{R}^{N \times T}$ denote the respective *sparse-change* components of $\mathbf{P}$ and $\mathbf{Q}$ corresponding to infrequent appliance activities. These two matrices will have piece-wise constant rows and the respective consecutive time differences of the entries, denoted by matrices $\mathbf{D}^P$ and $\mathbf{D}^Q$, are sparse. In other words, $\mathbf{S}^P$ becomes sparse under the linear transformation $\mathbf{S}^P = \mathbf{D}^P\mathbf{U}$ where $\mathbf{U} \in \mathbb{R}^{T \times T}$ is the upper triangular matrix of all ones; and similarly for $\mathbf{S}^Q = \mathbf{D}^Q\mathbf{U}$. This way, the number of non-zero entries $D_{n,t}^P$ and $D_{n,t}^Q$, that indicate the start/end of appliance activities, is much less compared to the total number of entries. Furthermore, each pair $\{D_{n,t}^P, D_{n,t}^Q\}$ is *jointly sparse*, as any appliance activities could manifest in *synchronized* changes of active/reactive power demand. This joint sparsity will be explored later for the load recovery problem.

Different from residential appliances, DERs like rooftop PVs would demonstrate correlated behaviors among nearby locations. Typically, PV generation output depends on localized factors such as solar irradiance and weather conditions [31]. Hence, the active power generated by rooftop PVs

within a neighborhood would share the same temporal pattern. This can be observed in Fig. 2 where half of the houses have PV installation and share a negative day-time component. We denote this underlying spatial correlation in $\mathbf{P}$ by a low-rank component $\mathbf{L} \in \mathbb{R}^{N \times T}$ of highly correlated rows. It will be further considered as a rank-one component later on. It is worth mentioning that the proposed low-rank plus sparse change model includes typical residential appliances and thus it encompasses general types of electricity consumption or generation in residential loads. For the interest of grid-edge monitoring, it will be mainly used for demonstrating the applications to EV event identification and BTM solar PV recovery later on.

Remark 2 (Reactive power support): Reactive power is assumed to have no spatial correlation which follows from the unity power factor setting under the IEEE Standard 1547-2018 [33]. This is currently the most common DER operation mode which does not provide grid reactive power support. Nonetheless, certain reactive power functions from DERs can be included, such as constant (non-unity) power factor or general Watt-Var curve where $\mathbf{Q}$ would share the same temporal pattern as $\mathbf{P}$. Hence, for general DER modes, we can slightly modify the data model to have the low-rank component also present in $\mathbf{Q}$.

Under the two load characteristics, the active power matrix can be decomposed into $\mathbf{P} = \mathbf{L} + \mathbf{D}^P \mathbf{U}$. This representation mimics the one in robust principal component analysis (RPCA), namely a low-rank plus sparse model [23]. To better reveal this model, we use the linear transformation that essentially computes the differences between consecutive columns (time instances) and obtain $\mathbf{P} = (\mathbf{K} + \mathbf{D}^P)\mathbf{U}$ where $\mathbf{K} := \mathbf{L}\mathbf{U}^{-1}$ is also of low rank. This is because of the temporal pattern in $\mathbf{L}$ leads to strong correlation of the differences as well. Similarly, we have $\mathbf{Q} = \mathbf{D}^Q \mathbf{U}$ which could include a low-rank component (see Remark 2). Clearly, the measurements in (1) and (6) are still linearly related to the unknown matrix components, namely $\mathbf{K}$, $\mathbf{D}^P$ and $\mathbf{D}^Q$.

The problem now boils down to recovering the low-rank and sparse matrix components, for which useful regularization terms can be introduced. The nuclear norm regularization is widely adopted in low-rank matrix recovery such as RPCA [23], subspace learning [24], and collaborative filtering [25]. It is defined as the sum of matrix's singular values, given by

$$\|\mathbf{K}\|_* := \sum_{i=1}^{N} \sigma_i(\mathbf{K}) \tag{7}$$

where $\sigma_i(\cdot)$ denotes the i-th largest singular value. The nuclear norm is a convex norm since it is the dual function of the matrix spectral norm; see [34, p. 637]. As for the sparse components, the popular L1-norm regularization defined as the sum of entry-wise absolute values is useful. The L1-norm is a tight convex relaxation of the L0 pseudo-norm (the total number of nonzero entries), and it has been shown to efficiently attain the sparse signal representation as widely used in compressed sensing and sparse signal recovery [35], [36]. Entries of $\mathbf{D} := [\mathbf{D}^P; \mathbf{D}^Q]$ are not only sparse but also *jointly*

sparse, because each appliance activity induces simultaneous changes in $\{D_{n,t}^P, D_{n,t}^Q\}$. Accordingly, we use the following *group* L1-norm definition:

$$\|\mathbf{D}\|_g := \sum_{n=1}^{N} \sum_{t=1}^{T} \left\| \left[D_{n,t}^P; D_{n,t}^Q \right] \right\|_2. \tag{8}$$

This group L1-norm can be thought of as the L1-norm for the jointly sparse pairs, and has been popularly used to promote sparsity at the group level [22]. Note that it is also convex as each summand term is the convex L2-norm.

Based on the two norm terms, we can formulate the following recovery problem:

$$\min_{\mathbf{K}, \mathbf{D}} \quad \|\mathbf{K}\|_* + \lambda \|\mathbf{D}\|_g \tag{9a}$$

$$\text{s. to} \quad -\mathcal{E} \leq \mathbf{\Gamma} - \begin{bmatrix} \mathbf{K} + \mathbf{D}^P \\ \mathbf{D}^Q \end{bmatrix} \mathbf{U}\mathbf{A} \leq \mathcal{E} \tag{9b}$$

$$-\mathbf{E} \leq \mathbf{Z} - \mathbf{H} \begin{bmatrix} \mathbf{K} + \mathbf{D}^P \\ \mathbf{D}^Q \end{bmatrix} \mathbf{U} \leq \mathbf{E} \tag{9c}$$

where the coefficient $\lambda > 0$ can balance between the two norms, while $\mathcal{E}$ and $\mathbf{E}$ contain the pre-determined entry-wise error bounds for the measurements. Measurement error can be bounded based on the metering accuracy or the model mismatch. If only measurement noise is considered, the bounds are set according to the largest metering error level for each datum. As Remark 1 points out, the linearized model (6) for $\mathbf{Z}$ would suffer from approximation error, which is included by the maximum deviation bounds in $\mathbf{E}$. Other types of error metrics may be used as well, such as the Frobenius norm which is preferred for Gaussian distributed error terms. For simplicity, the entry-wise maximum error bounds (9b)-(9c) have been used as linear constraints. Additional physics-based constraints such as non-negative $\mathbf{Q}$ may be considered as well.

Remark 3 (Choice of λ): To select λ, earlier work [23], [24] suggests to scale it with the matrix dimension as $\mathcal{O}(1/\sqrt{T})$ for $T \gg N$. We have followed this scaling with additional parameter tune-up. As mentioned earlier, this hyperparameter balances between the two matrix norms. Therefore, the value of λ would control the sparsity level of the recovered $\mathbf{D}$, or equivalently, the frequency of load changes. For very large λ, the recovered load profiles will be entirely contained in $\mathbf{L}$ with all-zero $\mathbf{D}$, and thus all loads are completely correlated with minimal individual activities. Similarly, under very small λ, there will be low correlation among the recovered profiles. Without the actual profiles available for validation, we have chosen the value of λ according to the typical occurrence of load change activities after adopting the $\mathcal{O}(1/\sqrt{T})$ scaling. Our numerical experiences suggest that the recovery performance is pretty robust to the λ choice. Thus, the parameter tuning step by matching the frequency of load changes can attain a satisfactory performance for the load recovery problem.

Since both norms are convex functions, (9) is a convex problem. The computational time associated with the nuclear norm, however, grows very quickly with the problem dimension as the resultant semidefinite program having a cubic computational complexity and a quadratic memory requirement [37]. To tackle this issue, we can simplify problem (9)

by assuming that the (relative) solar capacity is known for each house. This information can be obtained from customer reporting or historical data analysis by solving (9). Hence, the solar PV output is represented using a rank-one matrix as $\mathbf{K} = \mathbf{u}\mathbf{v}^{\mathsf{T}}$, where each entry of $\mathbf{u}$ is proportional to the solar capacity per house. This way, the problem now becomes a linear estimation one for $\mathbf{v}$ and $\mathbf{D}$, as given by

$$\min_{\mathbf{v},\mathbf{D}} \quad \frac{1}{2}\|\mathbf{v}\|_2^2 + \lambda \|\mathbf{D}\|_g \tag{10a}$$

$$\text{s. to} \quad -\boldsymbol{\mathcal{E}} \le \boldsymbol{\Gamma} - \begin{bmatrix} \mathbf{u}\mathbf{v}^{\mathsf{T}} + \mathbf{D}^P \\ \mathbf{D}^Q \end{bmatrix} \mathbf{U}\mathbf{A} \le \boldsymbol{\mathcal{E}} \tag{10b}$$

$$-\mathbf{E} \le \mathbf{Z} - \mathbf{H}\begin{bmatrix} \mathbf{u}\mathbf{v}^{\mathsf{T}} + \mathbf{D}^P \\ \mathbf{D}^Q \end{bmatrix} \mathbf{U} \le \mathbf{E} \tag{10c}$$

where the squared L2-norm of vector $\mathbf{v}$ is used to replace the nuclear norm regularization. This change follows from an alternative characterization of the nuclear norm of a rank-one matrix [25]. We will implement this simplified formulation for a large number of nodes.

Remark 4 (BTM solar disaggregation): The solar PV temporal pattern can be estimated by $\hat{\boldsymbol{\rho}}^T = \hat{\mathbf{v}}^{\mathsf{T}}\mathbf{U}$, where $\hat{\mathbf{v}}$ represents either the solution to (10) or the first right singular vector of the estimated $\mathbf{K}$ by (9). The estimated $\hat{\boldsymbol{\rho}}$ can be used to recover the total BTM solar output from the aggregated measurement $\mathbf{Z}^P$. To this end, the aggregated measurement per phase ϕ can be modeled as $\mathbf{z}_\phi^P \cong (\alpha\mathbf{1} + \beta\hat{\boldsymbol{\rho}}) + \mathbf{U}^{\mathsf{T}}\mathbf{d}_\phi^P$, where the unknown coefficients α and β transforms the solar pattern to the actual PV output while vector $\mathbf{d}_\phi^P$ is the total sparse load change for phase ϕ. Given $\hat{\boldsymbol{\rho}}$ and solar outputs being zero at non-daytime hours, one can use this linear model of $\mathbf{z}_\phi^P$ to estimate the three unknowns. The estimated $(\alpha\mathbf{1} + \beta\hat{\boldsymbol{\rho}})$ can separate the total solar outputs on phase ϕ.

Remark 5 (Periodic appliances): Appliance activities are assumed uncorrelated among houses, which may not hold for certain periodic loads such as HVACs. Residential HVAC consumption highly depends on local weather conditions, and thus its periodicity is really similar for co-located houses. As a result, the HVAC demand component can be thought of as the Fourier series expansions of a same periodicity, which is of low-rank as well. Theoretical results for the RPCA framework [23] imply that accurate recovery of the unknown matrix requires that the low-rank and sparse components do not overlap. Loosely speaking, the low-rank $\mathbf{K}$ is not sparse while the sparse $\mathbf{D}$ is not of low rank. Hence, the periodic appliance component, both sparse and low-rank, would be present in the recovered temporal pattern $\hat{\boldsymbol{\rho}}$, as demonstrated by numerical tests soon. A band-pass filter could resolve this issue by filtering out the signal of the HVAC periodicity (if known), which is out of scope for the present work.

IV. NUMERICAL TESTS

We have conducted numerical studies using real-world load datasets to evaluate the effectiveness of both recovery approaches (9) and (10). We have constructed two multiphase test systems by modifying the R2-12.47-3 feeder from the GridLab-D taxonomy feeder deposit [38]. We have one small 33-node feeder that can host 30 residential houses and

another 111-node one with 100 residential houses. In both systems, half of the houses are installed with solar PVs, and the minute-level active power matrix $\mathbf{P}$ has been obtained from the PecanStreet's Dataport [32] for homes located in Austin, TX. Active power data obtained for the months of December and July have been used to generate the winter and summer datasets, respectively. An example daily load profile for the 30 houses during a winter day is shown in Fig. 2. The reactive power matrix $\mathbf{Q}$ was formed by randomly selecting a power factor in the range [0.9, 0.95]. In addition, both $\mathbf{P}$ and $\mathbf{Q}$ were scaled to match the original loading of the R2-12.47-3 feeder.

To generate the D-PMU measurement data, we have run the nonlinear multiphase power flow simulation using GridLab-D [39] for both systems. Random noise was added to the simulated data to form the measurement $\mathbf{Z}$. We used the uniform distribution in the range of $[-0.02\%, 0.02\%]$ error for generating the noise in each D-PMU power measurement, per the total vector error of $\pm 0.01\%$ in a typical D-PMU datasheet [40]. The first D-PMU is always placed at the feeder head. Additional D-PMUs are placed at lateral heads to aggregate different subsets of down-stream loads. Intuitively, more D-PMUs can provide new information while the location of D-PMUs matters too. For example, D-PMUs installed at load nodes at the end of feeder are not as useful as those installed at the head of feeder or laterals. Hence, we have picked the D-PMU locations to aggregate as many load nodes as possible. As for the smart meter data, it was obtained by the averaging over every 15-minute window as in (1). We used the $\pm 0.2\%$ accuracy from the ANSI C12.20 Standard [41] to randomly generate uniformly-distributed additive noise for every smart meter measurement.

We have tested the original recovery Algorithm (9) using the 30-house system and the simplified one (10) using the 100-house system. Their performances in detecting night-time EV charging events and recovering the day-time BTM solar output have been evaluated. Actual load data during both winter and summer seasons have been considered to demonstrate the impact of large periodic appliances as discussed in Remark 5. To set up the optimization problem, the sparsity coefficient λ was fixed at 0.05 according to the $\mathcal{O}(1/\sqrt{T})$ scaling. The smart meter error bounds $\boldsymbol{\mathcal{E}}$ were determined using the aforementioned metering accuracy as 0.2% error for each entry. As for the D-PMU measurements, the error bounds $\mathbf{E}$ were set to be 0.2% per measurement base on analyzing the linear approximation error of (5). Moreover, the smart meter measurements for the 30-house system are assumed to *asynchronous*, while the 100-house system is assumed to be synchronous. Hence, we fix the measurement matrix $\mathbf{H}$ in (6) for the 30-house system according to the average loading. Meanwhile, this matrix has been updated using the latest smart meter data in the 100-house system (see Remark 1). Recovery Algorithms (9) and (10) are implemented using the CVX toolbox [42] with the MOSEK solver in the MATLAB R2020b simulator on a desktop with Intel Core i5 CPU @ 3.40 GHz and 16 GB of RAM.

A. Test Case 1 on 30-House Night-Time Data

We first demonstrate the capability of recovery Algorithm (9) in detecting EV events on the small feeder using the winter night-time dataset. We have increased the number of D-PMUs
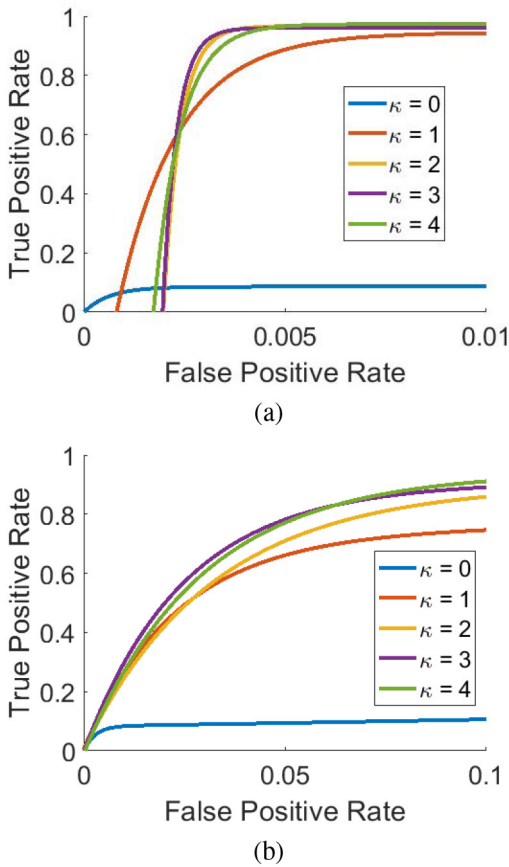
(a)

(b)

Fig. 3. Receiving operating curves (ROCs) of EV detection in Test Case 1 for the (a) winter and (b) summer night-time datasets under different number of D-PMUs.

(a)

(b)

Fig. 4. Recovered load data for one residential home with an EV charging event in Test Case 1 for the (a) winter and (b) summer night-time datasets.

from only one at the feeder head to a total of $\kappa = 4$ covering the full system. Recall that the additional D-PMUs are placed at lateral heads. The performance based on only smart-meter data is also included as denoted by $\kappa = 0$ (no D-PMUs). The EV start/end events are detected by selecting a threshold value as a percentage of the EV power rating, and a successful detection is declared only if both the actual house and time instance have been correctly identified. To evaluate the detection performance, the receiving operating curve (ROC) is plotted in Fig. 3(a) by comparing the false positive rate (FPR) versus the true positive rate (TPR) for the varying detection threshold. Thus, the closer the ROC is to the top left corner (TPR = 1 and FPR = 0), the better the detection performance is. To better visualize the ROCs, we have smoothed the curves using an exponential function fitting of all the points. For the case of $\kappa = 0$, Fig. 3(a) indicates that it is only possible to detect 8.7% of EV events with only smart meter data. This is exactly due to its low temporal resolution under the 15-minute averaging. By incorporating the total feeder power measurements using just one D-PMU, more than 94% of EV events can be correctly detected. Additional D-PMUs can further improve the TPR to detect almost 99% of EV events, thanks to the their high temporal resolution. For all D-PMU cases ($\kappa \geq 1$) the FPR is very small, since regular household appliances have much lower power ratings than EVs and it is unlikely to mistakenly declare their activities for an EV event. We have selected one residential house to illustrate the
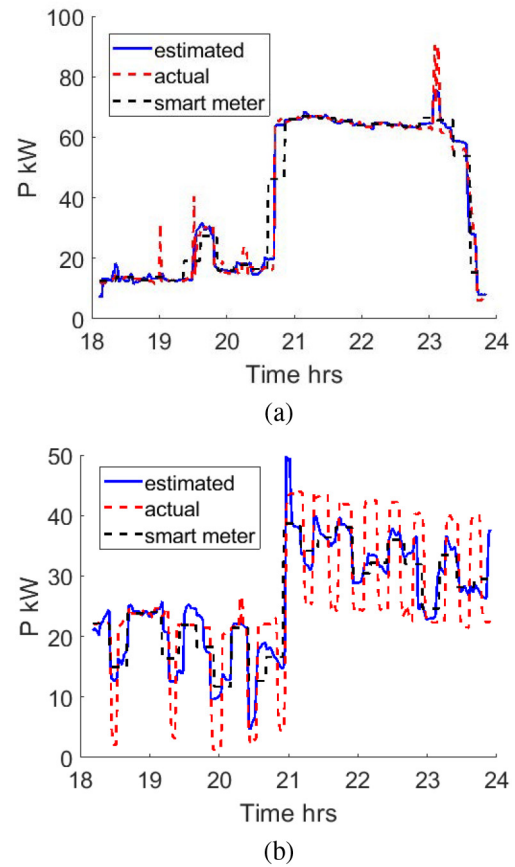
estimated profile from 4 D-PMUs during the night-time EV charging session, as shown in Fig. 4(a). Compared to the smart meter data, our estimated profile well matches the EV start/end charging time. Hence, by using just a few D-PMUs, recovery Algorithm (9) can effectively improve the spatial observability of the grid-edge DERs by identifying the exact location of EV typed activities in the feeder system.

We further test the small feeder using a summer night-time dataset, with the ROCs plotted in Fig. 3(b). As mentioned in Remark 5, HVAC loads in the summer violate the data modeling assumption for our proposed method, and thus have affected the detection performance. Compared to the winter results, the ROCs in Fig. 3(b) show reduced detection rates for all cases of κ. For one D-PMU, only 74.6% of EV events can be correctly identified, decreased from 94% earlier on. Additional D-PMUs have again enhanced the accuracy to over 80%, but all cases experience an increase of FPR to be around 0.1. This rising probability of false alarms is due to the high power rating of HVACs, which causes their activities to be declared as EV events by mistake. One estimated load profile during an EV start event (around 20:55) is also illustrated in Fig. 4(b). Compared to the winter recovery, the summer one again confirms the capability of identifying EV events, yet periodic HVAC activities are much more difficult to recover. Overall, our proposed method jointly utilizing D-PMU and smart meter data has attained enhanced spatial observability at the grid-edge.
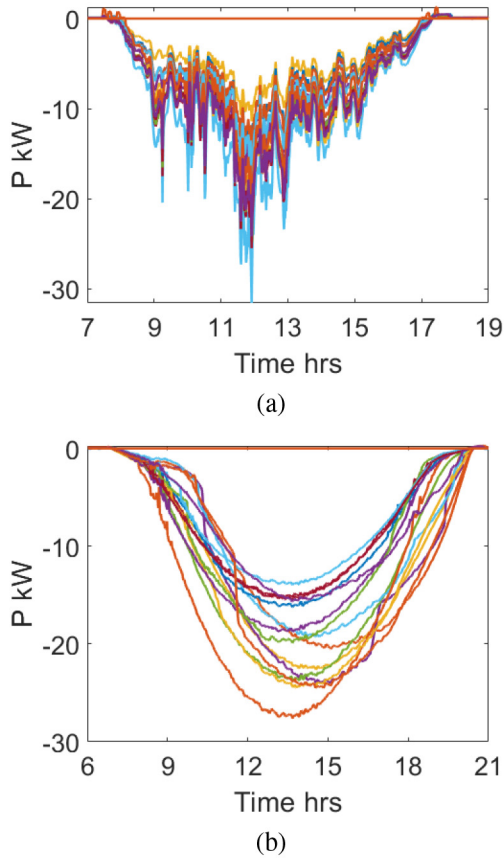
Fig. 5. Ground-truth PV outputs from (a) the winter and (b) summer datasets in Test Case 2 for all 30 houses.

Fig. 6. Recovered temporal patterns for the (a) winter and (b) summer day-time datasets in Test Case 2.

B. Test Case 2 on 30-House Day-Time Data

We also test recovery Algorithm (9) for recovering the BTM solar profile. The time window from hours 7:00-19:00 was used for the winter season and hours 6:00-21:00 for summer. These hours cover the entire day-time periods from sunrise to sunset. Figs. 5(a)–(b) plots the ground-truth PV outputs at all the houses for both datasets, respectively. The winter solar outputs exhibit high level of variability possibly due to cloud effects, while the summer ones tend to follow a smooth pattern of perfect solar irradiance. Due to the computer memory limitations, the recovery problem (9) has been solved for shorter time periods of 5-6 hours, with the (normalized) temporal patterns $\hat{\rho}$ recovered from the low-rank matrix solutions (as discussed in Remark 4) plotted in Fig. 6. The recovered winter solar pattern in Fig. 6(a) well matches the ground-truth profiles and even captures their fast variations around the noon hours for example. The summer solar pattern in Fig. 6(b) also well follows the parabolic trend of the ground-truth data, while showing an additional oscillation at a periodicity of 10-35 minutes. As discussed in Remark 5, the periodic HVAC activities would also affect the low-rank component, giving rise to this periodic behavior in the summer temporal pattern. To deal with this issue, we have used a simple band-pass filter to remove the frequency components in the periodicity range of 10-35 minutes from the estimated solar pattern (blue) to achieve a smoother pattern (red) as shown in Fig. 6(b). The recovered winter and filtered summer patterns have been effective for
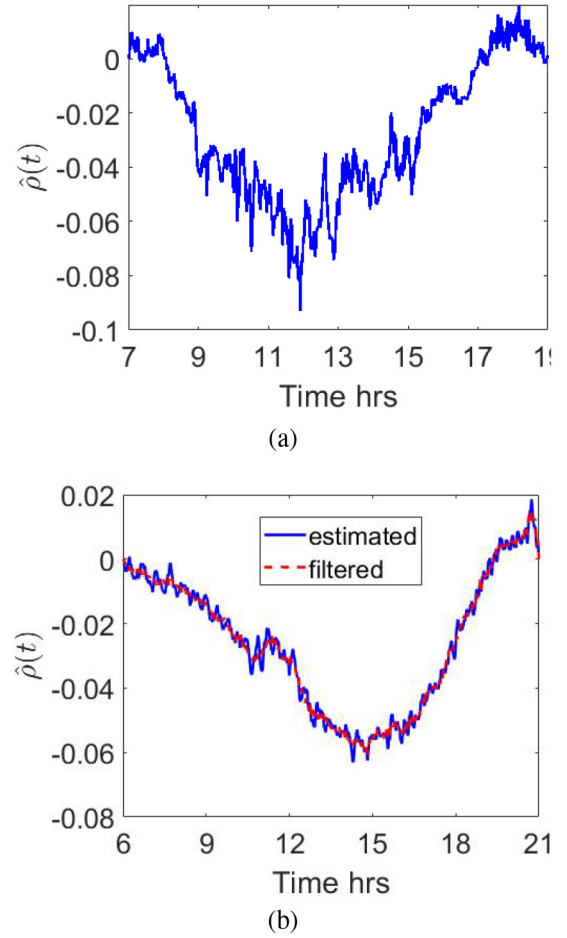
estimating the total solar output from the D-PMU aggregated power data, as shown by Fig. 7. Using the processing discussed in Remark 4, we scaled the temporal pattern to estimate the total BTM solar output from each aggregated power profile. While the winter solar estimation can perfectly recover the actual PV output, the summer one has some mismatch in the peak magnitude. Thanks to the capability of recovering the temporal pattern, the proposed method is useful for enhancing the temporal observability regarding the BTM solar generation. An important future direction will be utilizing historical data and direct solar PV measurements to improve the scaling of BTM solar estimation.

C. Test Case 3 on 100-House Night-Time Data

We demonstrate the computational improvement of the simplified method (10) by applying it to the 111-node feeder in this test case. The settings are similar to Test Case 1, and both winter and summer datasets are considered. The number of D-PMUs in the system has been varied from only one at the feeder head to a total of $\kappa = 7$. The ROCs comparing the FPR versus TPR are plotted in Figs. 8(a)–(b) for winter and summer respectively. Similarly to Test Case 1, the EV detection performance based on only smart meter data ($\kappa = 0$) is very poor due to its low temporal resolution.
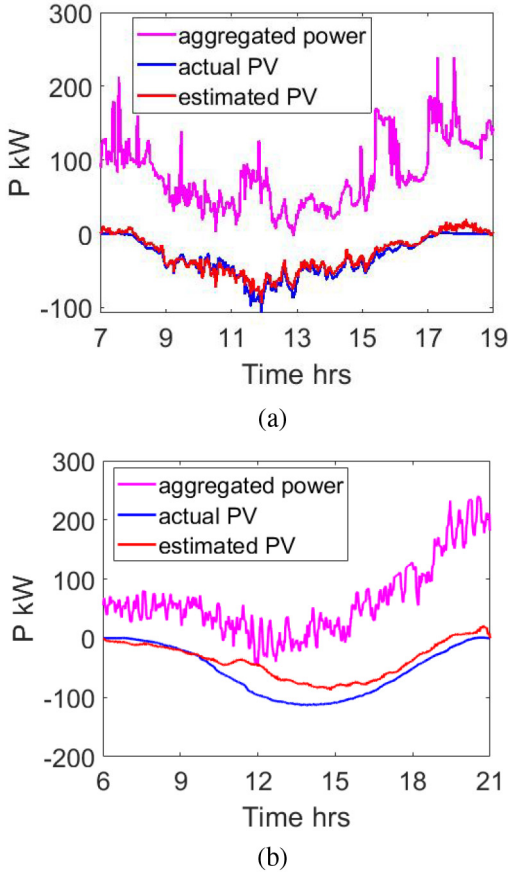
Fig. 7. Estimated solar PV outputs from the aggregated D-PMU power data for the (a) winter and (b) summer datasets in Test Case 2.

Fig. 8. ROCs of EV detection in Test Case 3 for the (a) winter and (b) summer night-time datasets with different number of D-PMUs.

Less than 10% of the EV events were correctly detected for both winter and summer datasets. By incorporating the D-PMU power data, Fig. 8(a) shows that the proposed method using (10) can increase the detection rates to $78.3\% - 95.5\%$ for the winter dataset. Compared to the small system results in Fig. 3(a), a larger number of D-PMUs is needed here to maintain the high identification accuracy due to the increase of house locations. Similar observations can be made for the summer EV detection shown in Fig. 8(b), where the maximum TPR ranging from $54\% - 73.4\%$ is lowered than the values in Fig. 3(b). Comparing between the two plots in Fig. 8 again reveals the effects of HVAC loads in increasing the likelihood of false alarms (larger FPR values) during the summer recovery. Overall, the FPR is relatively small as most household appliances have smaller power ratings than EVs. Therefore, the large system tests demonstrate our simplified method (10) can still effectively enhance the spatial observability regarding DER activities at the grid-edge.

D. Test Case 4 on 100-House Day-Time Data

We also used the simplified method (10) for recovering the BTM solar profile. The time windows and key steps all follow from Test Case 2. The procedure in Remark 4 was used to estimate the total BTM solar outputs as plotted in Fig. 9. A band-pass filter has again been implemented to smooth the recovered summer pattern in Fig. 9(b). Clearly, the recovered
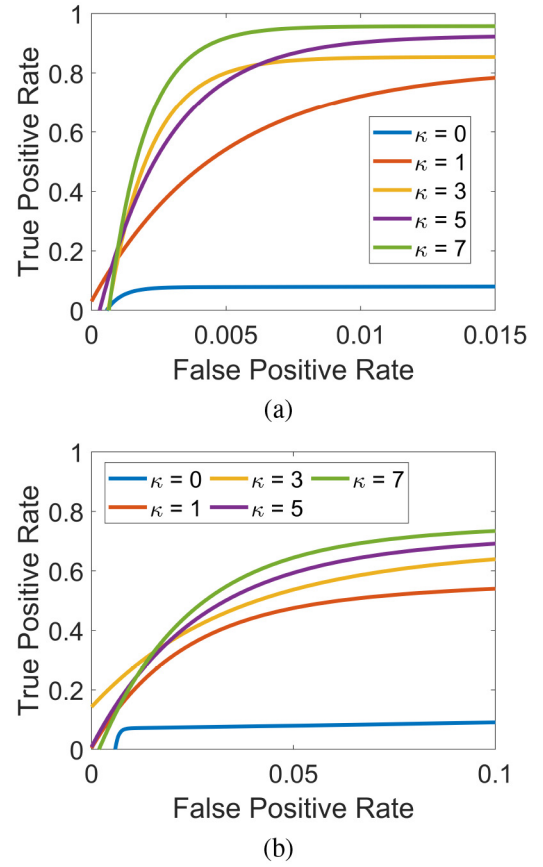
TABLE I
AVERAGE RUNTIMES OF (10) FOR VARYING NUMBER D-PMUS

κ	Runtime (seconds)
1	45.53
3	111.77
5	195.52
7	421.22

patterns well capture the trend of the actual solar PV outputs. Similar to Test Case 2, the recovered winter pattern can retain the fast solar transients while the summer one shows good match with the BTM PV output. Hence, the simplified method (10) using a rank-one component maintains the capability of effectively recovering the underlying temporal pattern in the residential loads. As a result, it has attained good accuracy in estimating the total BTM solar output from the aggregated power measurements in distribution systems.

Finally, Table I lists the average runtime for the simplified formulation (10) versus the number of D-PMUs κ for a 4-hour time period and the larger feeder. The runtime grows gradually with κ. Hence, additional D-PMUs can attain better monitoring performance (as shown by Test Case 3) at the cost of a slower runtime. Additionally, the runtime is affected by the length of load data due to the increasing problem dimension. We plan to develop adaptive online implementation in future to address this computational issue.
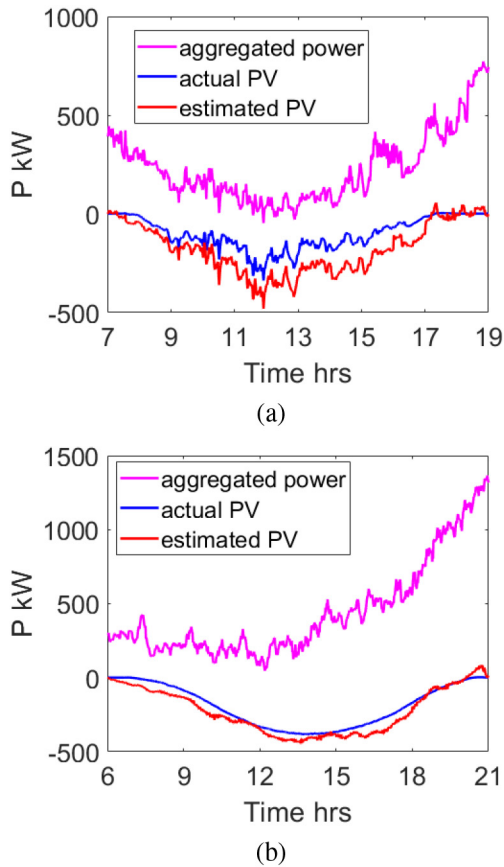
Fig. 9. Estimated solar PV outputs from the aggregated D-PMU power data for (a) winter and (b) summer datasets in Test Case 4.

V. Conclusion

This paper developed a spatio-temporal learning approach that can jointly harness the respective strengths of D-PMU and smart meter measurements to enhance the observability of DERs. To tackle the lack of measurements, we exploited the underlying characteristics of residential loads and connected DERs. Activities of EVs and other appliances are considered to be infrequent events, and a group L1-norm is used to enforce *jointly sparse* changes of active/reactive power. The strong correlation of solar generation within a feeder is modeled by a low-rank component, for which a nuclear norm is used. These two regularization terms have led to a convex recovery problem with linear constraints, which was further simplified by fixing the spatial scaling using the PV capacity information. Solutions to the recovery problems have been utilized to locate EV activities across the feeder and to infer the total BTM solar generation. Numerical tests using real-world load data on representative feeder systems have demonstrated the effectiveness of the proposed methods for monitoring DERs, although performance degradation has been observed in the presence of periodic HVAC loads.

Interesting future research directions open up for this work. We are currently working on improving the solar pattern recovery by utilizing additional data sources such as direct solar metering. Furthermore, we plan to investigate efficient algorithms for accelerated computation time and to incorporate the identified load characteristics to the development of modeling and control approaches for distribution systems.

References

[1] C. Niddodi, S. Lin, S. Mohan, and H. Zhu, "Secure integration of electric vehicles with the power grid," in *Proc. IEEE Int. Conf. Commun. Control Comput. Technol. Smart Grids (SmartGridComm)*, 2019, pp. 1–7.

[2] N. Jacobs *et al.*, "Analysis of system and interoperability impact from securing communications for distributed energy resources," in *Proc. IEEE Power Energy Conf. Illinois (PECI)*, 2019, pp. 1–8.

[3] F. Kabir, N. Yu, W. Yao, R. Yang, and Y. Zhang, "Joint estimation of behind-the-meter solar generation in a community," *IEEE Trans. Sustain. Energy*, vol. 12, no. 1, pp. 682–694, Jan. 2021.

[4] G. S. Ledva and J. L. Mathieu, "Separating feeder demand into components using substation, feeder, and smart meter measurements," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3280–3290, Jul. 2020.

[5] A. Von Meier, D. Culler, A. McEachern, and R. Arghandeh, "Microsynchrophasors for distribution systems," in *Proc. ISGT*, 2014, pp. 1–5.

[6] U.S. Department of Energy. (Sep. 2016). *Advanced Metering Infrastructure and Customer Systems*. [Online]. Available: https://www.energy.gov/sites/prod/files/2016/12/f34/AMI

[7] C. N. Lu, J. H. Teng, and W.-H. E. Liu, "Distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 10, no. 1, pp. 229–240, Feb. 1995.

[8] M. E. Baran and A. W. Kelley, "A branch-current-based state estimation method for distribution systems," *IEEE Trans. Power Syst.*, vol. 10, no. 1, pp. 483–491, Feb. 1995.

[9] S. Bhela, V. Kekatos, and S. Veeramachaneni, "Enhancing observability in distribution grids using smart meter data," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 5953–5961, Nov. 2018.

[10] P. L. Donti, Y. Liu, A. J. Schmitt, A. Bernstein, R. Yang, and Y. Zhang, "Matrix completion for low-observability voltage estimation," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2520–2530, May 2020.

[11] C. Carquex, C. Rosenberg, and K. Bhattacharya, "State estimation in power distribution systems based on ensemble Kalman filtering," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6600–6610, Nov. 2018.

[12] J. Zhao and L. Mili, "Robust unscented Kalman filter for power system dynamic state estimation with unknown noise statistics," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1215–1224, Mar. 2019.

[13] J. Song, E. Dall'Anese, A. Simonetto, and H. Zhu, "Dynamic distribution state estimation using synchrophasor data," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 821–831, Jan. 2020.

[14] G. Cavraro, E. Dall'Anese, and A. Bernstein, "Dynamic power network state estimation with asynchronous measurements," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, 2019, pp. 1–5.

[15] L. Schenato, G. Barchi, D. Macii, R. Arghandeh, K. Poolla, and A. Von Meier, "Bayesian linear state estimation using smart meters and PMUs measurements in distribution grids," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, 2014, pp. 572–577.

[16] A. Primadianto and C.-N. Lu, "A review on distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3875–3883, Sep. 2017.

[17] G. Wang, G. B. Giannakis, J. Chen, and J. Sun, "Distribution system state estimation: An overview of recent developments," *Front. Inf. Technol. Electron. Eng.*, vol. 20, no. 1, pp. 4–17, 2019.

[18] C. M. Cheung, W. Zhong, C. Xiong, A. Srivastava, R. Kannan, and V. K. Prasanna, "Behind-the-meter solar generation disaggregation using consumer mixture models," in *Proc. IEEE Int. Conf. Commun. Control Comput. Technol. Smart Grids (SmartGridComm)*, 2018, pp. 1–6.

[19] S. Lin and H. Zhu, "Enhancing the spatio-temporal observability of residential loads," in *Proc. 53rd Hawaii Int. Conf. Syst. Sci.*, 2020, pp. 1–10.

[20] N. Batra, Y. Jia, H. Wang, and K. Whitehouse, "Transferring decomposed tensors for scalable energy breakdown across regions," in *Proc. 32nd AAAI Conf. Artif. Intell.*, 2018, pp. 740–747.

[21] A. Miyasawa, Y. Fujimoto, and Y. Hayashi, "Energy disaggregation based on smart metering data via semi-binary nonnegative matrix factorization," *Energy Build.*, vol. 183, pp. 547–558, Jan. 2019.

[22] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *J. Royal Stat. Soc. B Stat. Methodol.*, vol. 68, no. 1, pp. 49–67, 2006.

[23] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, pp. 1–37, 2011.

[24] J. Wright, A. Ganesh, S. Rao, and Y. Ma, "Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization," Coordinated Sci. Lab., Univ. Illinois Urbana-Champaign, Rep. UILU-ENG-09-2210, DC-243, Urbana, IL, USA, 2009.

[25] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Rev.*, vol. 52, no. 3, pp. 471–501, 2010.

[26] W. H. Kersting, *Distribution System Modeling and Analysis*, 2nd ed. Boca Ranton, FL, USA: CRC Press, 2007.

[27] A. Bernstein, C. Wang, E. Dall'Anese, J.-Y. L. Boudec, and C. Zhao, "Load-flow in multiphase distribution networks: Existence, uniqueness, non-singularity and linear models," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 5832–5843, Nov. 2018.

[28] M. Bazrafshan and N. Gatsis, "Convergence of the Z-bus method for three-phase distribution load-flow with ZIP loads," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 153–165, Jan. 2018.

[29] E. Proedrou, "A comprehensive review of residential electricity load profile models," *IEEE Access*, vol. 9, pp. 12114–12133, 2021.

[30] M. Pipattanasomporn, M. Kuzlu, S. Rahman, and Y. Teklu, "Load profiles of selected major household appliances and their demand response opportunities," *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 742–750, Mar. 2014.

[31] N. Sharma, P. Sharma, D. Irwin, and P. Shenoy, "Predicting solar generation from weather forecasts using machine learning," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, 2011, pp. 528–533.

[32] Pecan Street, Inc. *PecanStreet Dataport*. Accessed: Jan. 2, 2021. [Online]. Available: https://www.pecanstreet.org/dataport/

[33] *IEEE Standard for Interconnection and Interoperability of Distributed Energy Resources With Associated Electric Power Systems Interfaces*, IEEE Standard 1547-2018 (Revision IEEE Std 1547-2003), 2018.

[34] S. Boyd and L. Vandenberghe, *Convex Optimization*, 1st ed. New York, NY, USA: Cambridge Univ. Press, 2004.

[35] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[36] D. L. Donoho, "For most large underdetermined systems of linear equations of minimal l_1-norm solution is also the sparsest solution," *Commnun. Pure Appl. Math.*, vol. 59, no. 6, pp. 797–829, 2006.

[37] K.-C. Toh and S. Yun, "An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems," *Pac. J. Optim.*, vol. 6, no. 3, pp. 615–640, 2010.

[38] GridLAB-D Taxonomy Feeders. Accessed: Jun. 3, 2021. [Online]. Available: https://github.com/gridlab-d/Taxonomy_Feeders

[39] *GridLab-D*. Accessed: May 4, 2021. [Online]. Available: https://www.gridlabd.org/

[40] Powerside. (2020). *MicroPMU Data Sheet*. [Online]. Available: https://powerside.com/wp-content/uploads/2020/01/Powerside_microPMU_DataSheet_.pdf

[41] American National Standards Institute. (2015). *ANSI C12.20-2015—Electricity Meters—0.1, 0.2, and 0.5 Accuracy Classes*. [Online]. Available: https://blog.ansi.org/2017/05/ansi-c1220-2015-electricity-meters-accuracy-classes/

[42] M. Grant and S. Boyd. (Jan. 2014). *CVX: MATLAB Software for Disciplined Convex Programming, Version 2.2*. [Online]. Available: http://cvxr.com/cvx

Shanny Lin (Graduate Student Member, IEEE) received the B.S. degree in electrical engineering from Rensselaer Polytechnic Institute, Troy, NY, USA in 2018, and the M.S. degree from the University of Texas at Austin, Austin, TX, USA, in 2020, where she is currently pursuing the Ph.D. degree. Her research interests include physics-aware and risk-aware machine learning for power distribution system operations.

Hao Zhu (Senior Member, IEEE) received the B.S. degree from Tsinghua University in 2006, and the M.Sc. and Ph.D. degrees from the University of Minnesota in 2009 and 2012. She is an Assistant Professor of Electrical and Computer Engineering (ECE) with the University of Texas at Austin. From 2012 to 2017, she was a Postdoctoral Research Associate and then an Assistant Professor of ECE with the University of Illinois at Urbana–Champaign. Her research focus is on developing innovative algorithmic solutions for problems related to learning and optimization for future energy systems. Her current interest includes physics-aware and risk-aware machine learning for power system operations, and designing energy management system accounting for cyber-physical coupling. She is a recipient of the NSF CAREER Award and the Siebel Energy Institute Seed Grant Award, and also the Faculty Advisor for three Best Student Papers Awarded at the North American Power Symposium. She is currently a member of the IEEE Power and Energy Society Long Range Planning Committee.