Distribution System State Estimation Using PV Separation Strategy in LV Feeders With High Levels of Unmonitored PV Generation

Amin Mokaribolhassan, *Graduate Student Member, IEEE*, Ghavameddin Nourbakhsh, *Member, IEEE*, Gerard Ledwich, *Life Fellow, IEEE*, Ali Arefi, and Mehdi Shafiei

***Abstract*—Distribution system state estimation (DSSE) is a critical analysis tool for active distribution networks (DNs). Unlike weighted least squares techniques, which are static DSSE methods, the augmented complex Kalman filter (ACKF) is a novel technique that considers the system's dynamic behavior. Currently, most DNs integrate a large number of unmonitored residential photovoltaic (PV) generations. Existing unmeasured PV sources violate the white noise assumption in Kalman and least-squares-based estimators, causing the estimator to be biased. Because the one-step difference of aggregated customer demand is characterized as white noise, the suggested PV estimation technique based on the differencing strategy is used to decouple PV from the measured load. Using the specified contribution factors, the new online pseudo current injections are generated. In addition, the estimator's accuracy is improved by using a new PV-scaling-aided ACKF approach based on the PV separation strategy. For validation purposes, this method is applied to real DN case studies. This study also makes use of an actual dataset to illustrate the efficacy of the proposed technique. The proposed technique outperforms the existing snapshot and dynamic DSSE techniques, and significant improvements are achieved in terms of accuracy and computational cost.**

***Index Terms*—Distribution system state estimation (DSSE), high photovoltaic (PV) penetration level, Kalman filter (KF), least squares, visualizing unmonitored residential PV.**

NOMENCLATURE

k	Time step.
j	Bus number.
n	Total number of buses.
m	Total number of branches.
s	Total number of samples.
$X(k)$	System state matrix at time step k.
$\hat{X}(k)$	Estimated state at time step k.
$Y(k)$	Measurement matrix at time step k.
$A(k)$	State transition matrix at time step k.
$C(k)$	Observation matrix at time step k.
$M(k)$	State noise matrix at time step k.
$N(k)$	Measurement noise matrix at time step k.
Q	State noise covariance matrix.
R	Measurement noise covariance matrix.
$V(k)$	Bus voltage vector at time step k.
V_0	Slack bus voltage vector.
$v_n(k)$	Bus n voltage at time step k.
$\hat{v}_n(k)$	Bus n estimated voltage at time step k.
$Ii(k)$	Bus current injection vector at time step k.
$Ib(k)$	Branch current vector at time step k.
$Ii^D(k)$	Bus current injection vector corresponding to the demand at time step k.
$Ii^{\text{PV}}(k)$	Bus current injection vector corresponding to PV at time step k.
$ii_n(k)$	Bus n current injection at time step k.
$ii_n^D(k)$	Bus n current injection corresponding to demand at time step k.
$ib_m(k)$	Branch m current at time step k.
$ii_n^{\text{PV}}(k)$	Bus n current injection corresponding to PV at time step k.
$p^M(k)$	Total measured active power on the transformer side at time step k.
$Q^M(k)$	Total measured reactive power on the transformer side at time step k.
$p^{\text{tPV}}(k)$	Total PV generated active power on the transformer side at time step k.
$p^{tD}(k)$	Total demand active power on the transformer side at time step k.
$p^{\text{RPV}}(k)$	Reference PV generated active power at time step k.
β^{PV}	PV scaling factor.
$w(k)$	PV modeling error at time step k.
P^M	Historical active power measurement vector on the transformer side.
P^{RPV}	Historical active power measurement vector of the reference PV.
P^D	Historical active power measurement vector of the total demand on the transformer side.
W	PV modeling error vector.
$P^{M'}$	Derivative vector of the historical active power measurements on the transformer side.

Manuscript received 21 September 2021; revised 22 December 2021; accepted 24 February 2022. Date of publication 4 April 2022; date of current version 24 February 2023. *(Corresponding author: Amin Mokaribolhassan.)*

Amin Mokaribolhassan, Ghavameddin Nourbakhsh, and Gerard Ledwich are with the School of Electrical Engineering and Robotics, Queensland University of Technology, Brisbane, QLD 4072, Australia (e-mail: a.mokaribolhassan@qut.edu.au; g.nourbakhsh@qut.edu.au; g.ledwich@qut.edu.au).

Ali Arefi is with the Murdoch University, Murdoch, WA 6150, Australia (e-mail: a.arefi@murdoch.edu.au).

Mehdi Shafiei is with the Acciona Energy, Melbourne, VIC 3000, Australia (e-mail: mehdi.shafiei.7206@gmail.com).

Digital Object Identifier 10.1109/JSYST.2022.3155164

$P^{\mathrm{RPV}'}$	Derivative vector of the reference PV historical measurements.
$\hat{\beta}^{\mathrm{PV}}$	Estimated PV scaling factor.
$\hat{p}^{\mathrm{tPV}}(k)$	Estimated total PV generation on the transformer side at time step k.
$\hat{p}^{tD}(k)$	Estimated total demand on the transformer side at time step k.
lf_j	Load contribution factor of bus j.
h_j^p	Total customer daily usage in bus j.
sf_j	Solar contribution factor of bus j.
s_j^p	Total installed PV size in bus j.
$P^D_{\mathrm{pseudo}}(k)$	Pseudo bus active power injection vector corresponding to the demand at time step k.
$P^{\mathrm{PV}}_{\mathrm{pseudo}}(k)$	Pseudo bus active power injection vector corresponding to the PV at time step k.
$Q^D_{\mathrm{pseudo}}(k)$	Pseudo bus reactive power injection vector corresponding to the demand at time step k.
$Ii^{\mathrm{PV}}_{\mathrm{pseudo}}(k)$	Pseudo current injection vector corresponding to PV at time step k.
$Ii^D_{\mathrm{pseudo}}(k)$	Pseudo current injection vector corresponding to demand at time step k.
$Ib_{\mathrm{pseudo}}(k)$	Pseudo branch current vector at time step k.
$V_{\mathrm{pseudo}}(k)$	Pseudo bus voltage vector at time step k.
$\hat{X}^{a+}(k)$	Posterior estimated augmented state after receiving the new measurement at time step k.
$\hat{X}^{a-}(k)$	Prior estimated augmented state before receiving the new measurement at time step k.
$P^{a+}(k)$	Posterior augmented error covariance after receiving the new measurement at time step k.
$P^{a-}(k)$	Prior augmented error covariance before receiving the new measurement at time step k.
$\Sigma(k)$	Innovation covariance at time step k.

I. Introduction

STATE estimation (SE) algorithms are used in networks to identify the most probable state of the network such as bus voltage or branch current using the available measurements. Static state estimation (SSE) methods are used to estimate the constant states. The weighted least squares (WLS) method is a commonly used SSE method in transmission networks [1], [2]. The drawback of the SSE methods is to ignore the state transition and the estimations from the previous samples, which are important in networks with high rates of renewable energies [3]. Therefore, dynamic SE methods known as forecasting-aided state estimation (FASE) are used to consider system dynamics [4]. Kalman filter (KF)-based strategies are commonly used in FASE methods [5]–[7]. The KF is a recursive estimator, which considers the dynamic approach of the system, compared with the WLS as a snapshot estimator [8], [9]. When the system is nonlinear and can be well approximated by linearization, the extended Kalman filter (EKF) is a good choice for the SE [3]. However, the EKF has some drawbacks, such as difficulty in the Jacobian calculation, high computational costs, and not working in highly nonlinear systems. To solve these issues, an unscented Kalman filter (UKF) that approximates the probability distribution is used [10]. Distribution system state estimation (DSSE) methods based on the UKF need more data to generate sigma points. In general, KF methods are based on nonlinear measurement function formulations. Furthermore, the difficulties of these formulations have required the consideration of linearization approaches.

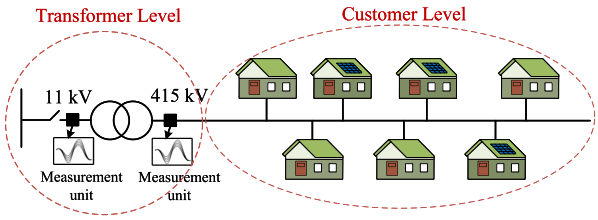

Fig. 1. Transformer and customer level in the distribution LV feeder.

DSSE methods that use a direct approach reduce the computational cost by linearizing the estimation formulation [11]. In the direct approach load flow solution, the constant bus-injection to branch-current (BIBC) and branch-current to bus-voltage (BCBV) matrices are used to solve the load flow problem. The relationship between system impedance, bus current injection, and node voltages is linear in the direct approach formulation. Therefore, the time-consuming procedures in Jacobian matrix calculation are not needed, and the ill-conditioned problems caused by lower–upper factorization of the Jacobian matrix will not occur in methods based on the direct approach. The accuracy of the load flow based on the direct approach is the same as other techniques, but it is both robust and efficient in large networks [11].

On the other hand, estimating the real and imaginary parts of the states independently is not sufficient due to the existing interactions between the real and imaginary parts. Therefore, the complex-valued states and the complex statistics are proposed in [8] and [12]. The augmented complex Kalman filter (ACKF) based on the direct approach is a noniterative DSSE technique that is used to overcome complex-valued nonlinear signals with strong levels of correlations. This method is applicable to three-phase distribution networks (DNs) due to the complicated nature of the power system. Recently, a new one-step multilayer DSSE based on the ACKF and the direct load flow approach is used to have a linear estimator in DNs with limited measurement units [12].

A. SE in DNs With Limited Measurement Devices

SE in DNs is challenging due to the limited available data. A distribution low-voltage (LV) feeder can be divided into transformer and customer levels shown in Fig. 1. In most of the DNs, measurement units on the transformer side are the only source of information that network operators have.

Installing phasor measurement units can improve the visibility of the network and linearize the measurement formulations in DSSE, not a cost-effective solution in current LV feeders [13]. Usually, indirect measurements called pseudo measurements are used to solve the estimation problem in networks with limited real measurements [14]. Pseudo measurements provide lower

accuracy and higher uncertainty compared with real measurements. Computation intelligence approaches, such as machine learning, can also be utilized to generate pseudo measurements in SE methodologies [15]. Learning-based methods require access to the ground-truth data, not applicable in DNs with limited measurements. Shafiei *et al.* [16], [17] consider the spatial–temporal dependencies among customer loads in DNs to improve the accuracy of the pseudo measurements and DSSE. The aggregation of adjacent loads has been used to decrease the variability of load consumption and improve the spatial correlation in the subarea to increase the accuracy of the pseudo measurements. According to [18] and [19], the accuracy of the DSSE in LV grids with limited measurements is mostly dependent on the accuracy of the pseudo measurements. On the other hand, high levels of photovoltaic (PV) penetration can deteriorate the accuracy of the generated pseudo measurements, which results in a biased estimator. The effect of PV penetration needs to be considered in pseudo measurement calculations, which have a great impact on the accuracy of the SE in DNs with limited measurements. According to [20], few papers have considered SE methods in DNs with high penetration of renewable resources. Ayiad *et al.* [21] present a study on the influence of PV penetration on the WLS-based DSSE. To increase the accuracy of the estimator, Barchi and Macii [22] use the available measured PV generation in the prediction and update steps of the DSSE. Measurement devices need to be installed at the customer level, which is not practical in most of the DNs with limited measurements.

B. Feeder-Level PV Estimation

Visualizing the unmonitored aggregated PV generation on the transformer side of the LV feeder can help the system operator to generate accurate pseudo data, which are used in the distribution state estimator. PV estimation techniques are widely used to estimate the individual customer's PV generation. These techniques are classified into data-driven and model-based methods. In model-based techniques, each customer's PV panel configuration data are required to estimate the solar output power, not practical in most of the feeders with limited data. Data-driven methods are more applicable than the model-based techniques in DNs with limited measurements [23], [24]. These methods are utilizing the available historical and online measurements to extract features to estimate the output power. Supervised, semisupervised, and contextually supervised data-driven approaches are used in literature to estimate the PV generation on the transformer or customer side [25]–[27]. Most of the data-driven techniques have the assumption of smart meters in each home for feature extraction and tuning [28]–[30]. This assumption is not practical in most of the DNs with unmonitored PV generations and limited measurement devices. Therefore, a new data-driven method is required to estimate the feeder-level PV generation in areas with limited measurement data. Since the aggregated demand has a noisy behavior in successive time steps, an unsupervised technique based on the differencing strategy can be utilized to visualize the unmonitored PV on the transformer side [31].

C. Contributions

The presence of unmonitored PV generations violates the white noise assumption in different Kalman and least-squares-based estimators. The white noise signal has independent random variables with Gaussian probability distribution and zero mean. On the other hand, PV generation in each bus would increase the correlation between each bus measurement (spatial correlation), which could violate the white noise assumption in the measurement noise. Additionally, PV generations can also increment the correlation in time between different samples (temporal correlation). Therefore, the white noise assumption for the state noise is also violated when PV exists in the network. A PV-scaling-aided ACKF (PV-ACKF) method is proposed in this article to increase the accuracy of the ACKF estimator in areas with high levels of PV penetration.

The contributions of the proposed method are also listed as follows:

1) developing the PV estimation technique based on the differencing strategy to be used in the state estimator;
2) providing a new strategy to generate pseudo measurements in DNs with limited measurement devices and high levels of PV penetration using the estimated PV scaling and contribution factors;
3) modifying the conventional ACKF formulation using the PV separation strategy and proposing the new PV-ACKF approach to be utilized in areas with high levels of PV penetration; and
4) evaluating different DSSE techniques in areas with high levels of PV penetration and limited measurement devices, which are not studied in the previous works.

The rest of this article is organized as follows. Section II describes the conventional ACKF. Section III presents the proposed method used in this study. Section IV discusses the simulation results. Finally, Section V concludes this article.

II. DSSE Based on the ACKF and the Direct Approach

A current injection-based estimator was previously proposed in [32] that converts the power injection measurements to the bus current injections in a rectangular form. The complex nature of the bus current injections as the system states necessitates a complex formulation of DSSE algorithms. While estimating the real and imaginary portions of the states separately reduces the algorithm's complexity, it overlooks the interactions between the real and imaginary parts [12]. On the other hand, DSSE based on a direct power flow approach can linearize the estimation formulation [11]. The ACKF based on the direct approach is a new FASE approach developed and applied in [8] and [33]. According to [8], the ACKF has consistently better performance in terms of computation and accuracy compared with traditional WLS.

The general state-space model is presented as follows:

$$\begin{aligned} X(k) &= A(k)X(k-1) + M(k) \\ Y(k) &= C(k)X(k) + N(k). \end{aligned} \tag{1}$$

Using (1), the augmented state-space model can be written as follows [16]:

$$\begin{aligned} X^a(k) &= A^a(k)X^a(k-1) + M^a(k) \\ Y^a(k) &= C^a(k)X^a(k) + N^a(k) \\ X^a(k) &= \begin{bmatrix} X(k) \\ X^*(k) \end{bmatrix}, \quad Y^a(k) = \begin{bmatrix} Y(k) \\ Y^*(k) \end{bmatrix} \\ A^a(k) &= \begin{bmatrix} A(k) & 0 \\ 0 & A^*(k) \end{bmatrix}, \quad C^a(k) = \begin{bmatrix} C(k) & 0 \\ 0 & C^*(k) \end{bmatrix} \\ M^a(k) &= \begin{bmatrix} M(k) \\ M^*(k) \end{bmatrix}, \quad N^a(k) = \begin{bmatrix} N(k) \\ N^*(k) \end{bmatrix} \end{aligned} \tag{2}$$

where subscripts $(..)^a$ and $(..)^*$ show the augmented complex and complex conjugate operator, respectively. Here, the noise properties need to be white. System and measurement noises need to be uncorrelated in time and have the property of Gaussian normal distribution with zero mean and specific variances. This basic assumption applies to all the DN SE techniques to have an unbiased estimator, which can be violated due to the existing uncertainty sources such as PV.

In this article, like [32], system states are bus current injections in complex form. Based on the study in [12], the $A^a(k)$ matrix is considered to be an identity in (2) since the difference between current injections in each bus for successive time steps is characterized as white noise. This assumption is valid when there are only customers connected to each bus and the level of PV penetration is low.

Using the direct load flow approach formulations, the measurement model can be obtained. Here, BCBV and BIBC matrices, which are obtained from the topological characteristics of the system, need to be used to solve the load flow problem offline. The multiplication of these two matrices provides the direct load flow (DLF) matrix, which indicates the linear relationship between bus current injection and the bus voltage to have a linear estimator. The measurement model can be given by the following equations at time step k:

$$\begin{aligned} Ib(k) &= \text{BIBC} \times Ii(k) \\ V(k) - V_0 &= -\text{DLF} \times Ii(k) \end{aligned} \tag{3}$$

$$Ib(k) = \begin{pmatrix} ib_1(k) \\ ib_2(k) \\ \vdots \\ ib_m(k) \end{pmatrix}, \quad Ii(k) = \begin{pmatrix} ii_1(k) \\ ii_2(k) \\ \vdots \\ ii_n(k) \end{pmatrix}$$

$$V(k) = \begin{pmatrix} v_1(k) \\ v_2(k) \\ \vdots \\ v_n(k) \end{pmatrix}, \quad V_0 = \begin{pmatrix} v_1(k) \\ v_1(k) \\ \vdots \\ v_1(k) \end{pmatrix} \tag{4}$$

where $Ib(k) \in \mathbb{C}^{m\times 1}$, $Ii(k) \in \mathbb{C}^{n\times 1}$, $V(k) \in \mathbb{C}^{n\times 1}$, $V_0 \in \mathbb{C}^{n\times 1}$, BIBC $\in \mathbb{R}^{m\times n}$, and DLF $\in \mathbb{C}^{n\times n}$. The standard KF is used in the ACKF formulation since the state and the measurement functions are linear. (The relationship between bus current injection and voltage is linear since direct approach is utilized.)

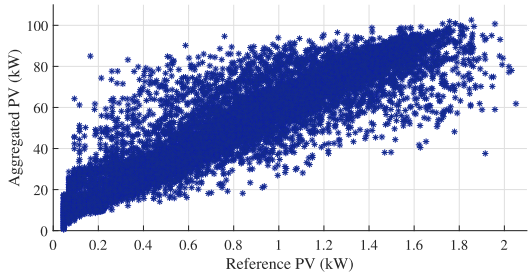

Fig. 2. Correlation between the RPV and the aggregated PV on the transformer side using the real dataset for one LV feeder.

The ACKF is characterized by rapidly converging to the true values [12].

III. Proposed Method

In this section, a new estimation method based on PV separation is proposed to solve the problems associated with the conventional ACKF in areas with a high level of PV penetration. Since PV measurements are not available, an unsupervised PV estimation technique is required to separate different portions from the aggregated measurement on the transformer side. Then, pseudo measurements are generated to make the problem of estimation solvable. Finally, PV-ACKF formulation is suggested to provide the new FASE technique to be used in areas with a high level of PV penetration and limited measurements.

A. PV Estimation Technique

Here, the new unsupervised data-driven technique is used to identify the unknown portions of the measured aggregated values on the transformer side. In this study like [12], active and reactive power measurements are available on the transformer side, which is common in most of the DNs.

1) Correlation Studies and PV Scaling Factor: Since all customers are on the same LV feeder, there is a correlation between their PV generation outputs. According to [29] and [31], one of the customers' PV output can be used as a reference PV (RPV) to estimate the total unmonitored PV generation on the transformer side due to the existing correlation between the RPV and the aggregated PV output on the transformer side. Fig. 2 shows the correlation between the RPV and the aggregated PV on the transformer side using a one-month real dataset with a resolution of 1 min for an area in the range of 1 km.

Radiance measurements can also be used to generate the RPV data, which need additional measurement devices, such as sky cameras, to monitor clouds and also panel configuration information to have a correct model. In this study, RPV solar output is monitored to estimate total PV on the transformer side.

The linear correlation can be assumed between the RPV's generated active power and the total PV generation on the transformer side of the LV feeder [29]. However, nonlinearity between the RPV measurements and the total PV could also occur, caused by transposition errors. The nonlinear model is not a viable option since it necessitates extra high-resolution measurements to solve these issues, which are inapplicable

in present DNs with restricted measurement capabilities [34]. Equation (5) can be expressed for any time k based on the linear relationship between RPV active power generation and the aggregated total PV active power on the transformer side

$$p^{\text{tPV}}(k) = \beta^{\text{PV}} \times p^{\text{RPV}}(k) + w(k). \tag{5}$$

PV modeling error ($w(k)$) can have a characteristic similar to white noise. White noise has a Gaussian distribution with a mean of zero and no temporal correlation [31]. This assumption is achievable if all of the PV panels are connected to a single LV feeder.

The PV scaling factor (β^{PV}) can be updated whenever the network gives updated data on new PV installation, PV panel deterioration, shadowing due to environmental changes, or seasonality effects [31].

2) PV Scaling Factor Estimation: Since current and voltage phasor measurements are available on the transformer side of the LV feeder, the measured total active power ($p^M(k)$) can be obtained using power calculations. The measured total active power consists of unknown PV and demand parts, as follows:

$$p^M(k) = p^{tD}(k) - p^{\text{tPV}}(k). \tag{6}$$

Then, (7) and (8) are obtained, which take into account two known vectors (P^M and P^{RPV}) and one unknown aggregated demand vector (P^D):

$$P^M = P^D - \beta^{\text{PV}} \times P^{\text{RPV}} + W \tag{7}$$

$$P^M = \begin{pmatrix} p^M(1) \\ p^M(2) \\ \vdots \\ p^M(k) \end{pmatrix}, \quad P^D = \begin{pmatrix} p^{tD}(1) \\ p^{tD}(2) \\ \vdots \\ p^{tD}(k) \end{pmatrix},$$

$$P^{\text{RPV}} = \begin{pmatrix} p^{\text{RPV}}(1) \\ p^{\text{RPV}}(2) \\ \vdots \\ p^{\text{RPV}}(k) \end{pmatrix}, \quad W = \begin{pmatrix} w(1) \\ w(2) \\ \vdots \\ w(k) \end{pmatrix}. \tag{8}$$

Because there are correlations between P^D and P^{RPV} data streams, estimating the PV scaling factor just from the measured P^M vector would be a difficult process. In this article, signal processing analyses are used to find β^{PV}. According to [12], the first-order difference of the aggregated customer demand vector (P^D) acts like white noise, which will aid to remove the unmeasured part from the formulation. Therefore, the derivative operator is utilized and (7) is updated to

$$D\left(P^M\right) = D\left(P^D\right) - D\left(\beta^{\text{PV}} \times P^{\text{RPV}}\right) + D(W). \tag{9}$$

Here, $D(..)$ represents the first-order derivative operator. P^M and P^{RPV} are the historical measurement vectors. Since $D(P^D)$ acts like white noise, the least squares solution provides the optimal β^{PV} using the following [31]:

$$\hat{\beta}^{\text{PV}} = \left([P^{M'}]^T [P^{\text{RPV}'}]\right)^{-1} [P^{\text{RPV}'}]^T [P^{M'}]. \tag{10}$$

Using the estimated scaling factor (10) and the new measurement of the active power on the transformer side, the online estimated demand and PV portions on the transformer side are obtained using (11). These estimated values are utilized in the next stage to generate the real-time pseudo measurement to be used in the SE formulation

$$\hat{p}^{\text{tPV}}(k) = \hat{\beta}_{\text{PV}} \times p^{\text{RPV}}(k)$$
$$\hat{p}^{tD}(k) = p^M(k) - \hat{\beta}_{\text{PV}} \times p^{\text{RPV}}(k). \tag{11}$$

B. Pseudo Measurements

Since the measurements in an LV feeder are mostly limited to the transformer side, pseudo measurements are needed to use in the SE formulation. Accuracy of the pseudo measurements plays an important role in improving the accuracy of the DSSE in networks with limited measurements. However, historical measurements cannot be efficient to generate real-time pseudo measurements in the SE. Here, pseudo measurements are generated and updated using the measurements on the transformer side. Like [12], pseudo measurements are generated using the customer's average daily usage (kWh) (available in the electricity bill) as historical data to find each bus's load contribution factor. The load contribution factor for bus j is obtained as

$$lf_j = \frac{h_j^p}{\sum_{j=1}^{n} h_j^p} \tag{12}$$

where h_j^p is the total customer daily usage at bus j using the billing information. The load factor for each bus j is constant for different time steps since the geographical aggregation of individual customer consumption has a smoothing effect [12]. The real-time pseudo bus active power corresponding to demand is obtained using the calculated load contribution factor vector LF and the estimated demand part of the measured active power on the transformer side (13)

$$P^D_{\text{pseudo}}(k) = LF \times \hat{p}^{tD}(k) \tag{13}$$

where $P^D_{\text{pseudo}}(k) \in \mathbb{R}^{n\times 1}$ and $LF = [lf_1, lf_2, \ldots, lf_n]^T$. Having PV generations in the LV feeder would violate the assumption of white noise in the ACKF formulation, whereby generated pseudo measurements could not be accurate. Consequently, the conventional ACKF is biased in networks with high PV penetration levels. To mitigate this issue, it is imperative to design a new formulation to generate pseudo measurements.

In the conventional ACKF, pseudo measurements are generated using the load contribution factors. In this article, the PV contribution factor is proposed to improve the accuracy of the generated pseudo measurements in areas with unmonitored PV. In general, the DN operator has access to the registered PV sizes of each customer. The contribution of bus j to the estimated total PV active power on the transformer side is calculated as follows:

$$sf_j = \frac{s_j^p}{\sum_{j=1}^{n} s_j^p}. \tag{14}$$

Using the total installed PV size at each bus (s_j^p) divided by the total PV size in the feeder, the PV contribution factor is calculated. Using the obtained PV contribution factor (sf_j), the real-time pseudo bus active power corresponding to the PV is obtained as

$$P^{\text{PV}}_{\text{pseudo}}(k) = SF \times \hat{p}^{\text{tPV}}(k) \tag{15}$$

where $P^{\text{PV}}_{\text{pseudo}}(k) \in \mathbb{R}^{n\times 1}$ and $SF = [sf_1, sf_2, \ldots, sf_n]^T$. Furthermore, the pseudo reactive power in each bus is also obtained using the load contribution factor and the total reactive power (Q^M) in (16). The measured total reactive power is calculated using the measured current and voltage on the transformer side. PV inverters are assumed not to inject reactive power into the network. In scenarios that solar inverters are injecting reactive power, network operators need to define their power factor or install monitoring devices

$$Q^D_{\text{pseudo}}(k) = LF \times Q^M(k). \tag{16}$$

Using the direct power flow approach (see [11]), the pseudo current injection vector corresponding to PV ($Ii^{\text{PV}}_{\text{pseudo}}(k)$) and demand ($Ii^D_{\text{pseudo}}(k)$) portions can be obtained. Furthermore, pseudo branch currents ($Ib_{\text{pseudo}}(k)$) and bus voltage ($V_{\text{pseudo}}(k)$) are found to be used in the estimator.

C. PV-Scaling-Aided ACKF

In the SE procedure, state and measurement noises are required to have a white noise characteristic. Existing PV generation would violate this basic assumption since there are correlations between the PV outputs in successive time steps and measurements in different buses due to the similar characteristic between solar outputs of customers in one feeder. On the other hand, the uncertainty of the cloud movements can also make the estimation process challenging. The bus current injection can be disaggregated to the corresponding demand and PV portions in each bus using the following:

$$Ii(k) = Ii^D(k) - Ii^{\text{PV}}(k) \tag{17}$$

$$Ii^D(k) = \begin{pmatrix} ii^D_1(k) \\ ii^D_2(k) \\ \vdots \\ ii^D_n(k) \end{pmatrix}, \quad Ii^{\text{PV}}(k) = \begin{pmatrix} ii^{\text{PV}}_1(k) \\ ii^{\text{PV}}_2(k) \\ \vdots \\ ii^{\text{PV}}_n(k) \end{pmatrix}. \tag{18}$$

Here, $Ii^D(k) \in \mathbb{C}^{n\times 1}$ and $Ii^{\text{PV}}(k) \in \mathbb{R}^{n\times 1}$. In this article, the system model needs to be redesigned to meet the white noise assumption. In the updated system model, the system state is considered to be the demand portion of the bus current injection ($Ii^D(k)$). Here, the A matrix is considered to be identified since the difference between bus current injections corresponding to demand for successive time steps follows white noise characteristics [33].

The proposed flowchart for the PV-ACKF is presented in Fig. 3 where the measurement model is updated as

$$\underbrace{\begin{bmatrix} Ii^D(k) \\ Ib(k) \\ V(k) - V_0 \end{bmatrix}}_{Y(k)} = \underbrace{\begin{bmatrix} I & 0 & 0 \\ 0 & \text{BIBC} & 0 \\ 0 & 0 & -\text{DLF} \end{bmatrix}}_{C} \times \underbrace{\begin{bmatrix} Ii^D(k) \\ Ii^D(k) - Ii^{\text{PV}}_{\text{Pseudo}}(k) \\ Ii^D(k) - Ii^{\text{PV}}_{\text{Pseudo}}(k) \end{bmatrix}}_{X(k)}. \tag{19}$$

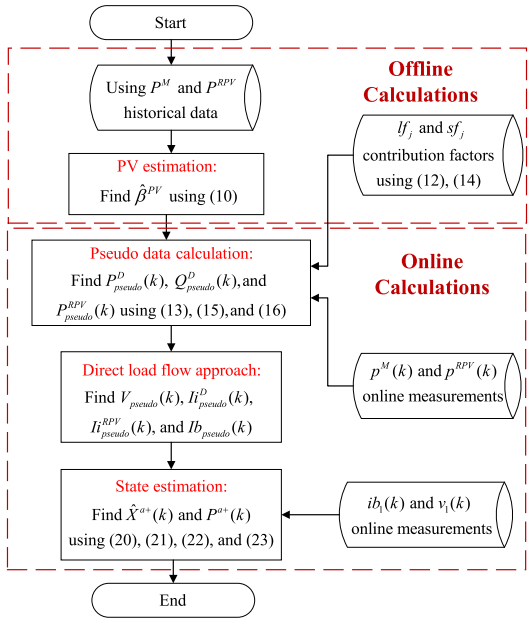

Fig. 3. Proposed PV-ACKF flowchart.

In the first step, the PV scaling factor is estimated using the aggregated total active power on the transformer side and the RPV active power generation offline. Besides, PV and demand contribution factors are also found offline using the customer's billing information and registered PV sizes. Real-time pseudo measurements are generated at each time step k using the estimated values on the transformer side and the calculated contribution factors. Then, pseudo bus current injections corresponding to demand and PV portions are obtained using the direct power flow analysis. Finally, the PV-ACKF SE stage is providing the optimal estimate. The SE procedure is formulated using the following steps:

1) initializing the estimator using the augmented state ($\hat{X}^a(0)$) and the augmented error estimate covariance ($P^a(0)$) calculated as

$$\hat{X}^a(0) = E\left(X^a(0)\right)$$
$$P^a(0) = E\left[\left(X^a - \hat{X}^a(0)\right)\left(X^a - \hat{X}^a(0)\right)^T\right] \tag{20}$$

2) predicting the new state and error covariance using

$$\hat{X}^{a-}(k) = \hat{X}^a(k-1)$$
$$P^{a-}(k) = A^a(k)P^a(k-1)A^{aT}(k) + Q^a(k-1) \tag{21}$$

3) Kalman gain ($K(k)$) calculation using

$$K(k) = P^{a-}(k)C^{aT}(k)\Sigma^{-1}(k)$$
$$\Sigma(k) = C^a(k)P^{a-}(k)C^{aT}(k) + R^a(k) \tag{22}$$

4) state and error covariance update using

$$\hat{X}^{a+}(k) = \hat{X}^{a-}(k) + K(k)(Y^a(k) - C^a(k)\hat{X}^{a-}(k))$$
$$P^{a+}(k) = (I - K(k)C^a(k))P^{a-}(k). \tag{23}$$

In the prediction step, the previous time estimated augmented state ($\hat{X}^a(k-1)$) and the previous time augmented error covariance ($P^a(k-1)$) are replaced by the initial values calculated in (20). Then, an optimal Kalman gain needs to be calculated to put weight to the received measurements and predicted states to update the new state. The last step is the updating stage, where new states and error covariance are calculated using Kalman gain and measurement function. The obtained new state and error covariance are assumed as the previous state and be replaced in the prediction step to find the new state.

D. Outlier Detection

Outliers can appear in a variety of ways throughout the measurement process, causing system measurements to be contaminated. To have a robust estimation, outliers need to be detected and a new estimation mechanism should be developed. A typical practice across various approaches is to identify sufficient statistics that have information on outliers and then design a statistical hypothesis test to determine whether or not the present data sample contains an outlier. To detect outliers, the innovation quantity needs to be obtained from the update stage of the estimation. The innovation $r(k)$ can be written as

$$r(k) = Y^a(k) - C^a(k)\hat{X}^{a-}(k). \tag{24}$$

In an ideal situation, the innovation needs to follow a normal distribution with zero mean and a covariance, which equals $\Sigma(k)$ in (22) [5]. Using

$$d(k) = \sqrt{r^T(k)\Sigma^{-1}(k)r(k)} \tag{25}$$

the Mahalanobis distance ($d(k)$) is calculated to compare with a threshold value to verify outlier presence. Random variables need to be within the $3d(k)$ band, and the measured data out of this range can be considered as a bad data measurement. According to [12], to prevent omitting important information from the measured data due to the presence of PV and pseudo measurements, $5d(k)$ is considered as a threshold value to detect outliers. When the outlier is detected, the measurement needs to be discarded, and only the prediction step of the estimation procedure [see (21)] is used.

E. Evaluation Metrics

In this article, mean absolute error (MAE) and mean error variance (MEV) metrics are used to evaluate the performance of the proposed method. The MAE is computed for voltage magnitude and angle across all tests as follows [31]:

$$\text{MAE} = \frac{1}{s}\sum_{\text{k}=1}^{s} |v_n(k) - \hat{v}_n(k)|. \tag{26}$$

Besides, the MEV is used to determine the state estimator's actual error variance using the following equations [8]:

$$\text{MEV} = \frac{1}{s}\sum_{k=1}^{s} (e_n(k) - \bar{e}_n)^2 \tag{27}$$

$$ce_n(k) = v_n(k) - \hat{v}_n(k) \tag{28}$$

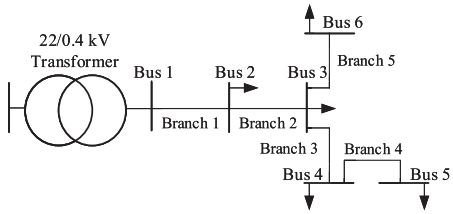

Fig. 4. Six-bus DN [12].

where $e_n(k)$ and $\bar{e}_n$ are the estimation error of bus n at time step k and the total average error of bus n, respectively.

IV. Simulation Results

In this section, simulation results using the proposed method are presented. The accuracy of the estimates using the proposed PV-ACKF is compared with the results of the conventional techniques. The real-world dataset is far more advantageous than the synthetic one. A synthetic dataset is highly dependent on the generating model, which needs to be perfect to mimic the authentic data. Modeling cloud movements and residential PV generations is difficult in areas with limited information.

In this article, the real dataset provided by Energex and Ergon Energy Network (the local DN service provider in Southeast Queensland) is used for validation purposes. The data were gathered by Redback measurement devices along the LV feeder and on the transformer side. This 1-min-resolution dataset is for residential customers located in one LV feeder for a two-month period from November 2020 to December 2020. Data exploration, cleaning, integration, and transformation were done to prepare the data for doing analysis.

The proposed PV-ACKF method is tested on three different case studies to show the effectiveness of the algorithm in different networks. Here, a real six-bus radial DN [12], an unbalanced 23-bus Australian DN [16], and a modified IEEE-116 LV test feeder [35], [36] are used. For all of the case studies, the measuring point is considered on the secondary side of the medium-voltage/LV transformer (available for most of the LV feeders).

The simulations are run on a computer with a 3.4-GHz Intel Core i7-4770 CPU and 16 GB of RAM. MATLAB R2018a software is used to estimate the PV generation, and Spyder (Python 3.8) software is utilized to do data preparation and run the SE techniques using MATLAB outputs.

A. Six-Bus DN

The six-bus DN presented in Fig. 4 is used in this study. This LV feeder incorporates 67 domestic homes, where 33 customers have rooftop PV systems. Here, customer groups are connected to each bus, and all the installed load and PV information are presented in Table I. The total installed PV capacity is 120.2 kWp (kilowatt peak), and the maximum total demand on the transformer side is 122.49 kW (kilowatt). The total PV penetration level for this network is 98% obtained using the total original installed PV capacity over the maximum demand. Here, installed

TABLE I
SIX-BUS-TEST-SYSTEM-CONNECTED LOAD AND PV INFORMATION

Bus	Num. of loads	PV size (kWp)	Max. demand (kW)	PV penetration (%)
2	10	14.4	22.91	56
3	15	30.2	42.61	71
4	15	28.3	33.46	84
5	12	24.7	26.1	95
6	15	22.6	42.86	53

TABLE II
LOAD AND PV CONTRIBUTION FACTORS IN PERCENT FOR DIFFERENT SCENARIOS

Bus	Scenario 1		Scenario 2		Scenario 3	
	LF(%)	SF(%)	LF(%)	SF(%)	LF(%)	SF(%)
2	11	12	19	27	11	0
3	27	25	28	57	20	0
4	22	23	22	16	22	16
5	15	21	12	0	19	27
6	25	19	19	0	28	57

measurement units are limited to substation levels. Only bus voltage and branch current measurements are available on the transformer side. Moreover, one of the customers' PV output is measured to be used in PV-ACKF formulation.

1) Generating Pseudo Measurements Using PV Separation Strategy: In the conventional ACKF, pseudo measurements are generated using the load contribution factors without considering the impact of PV generations. In this article, a new approach to generating accurate pseudo measurements using PV contribution factors is proposed. Here, different scenarios are considered to represent the impact of PV feeder distribution on the accuracy of the pseudo measurements. In the first scenario, the installed PV units are evenly distributed along the LV feeder. According to the second and third scenarios, most of the PV units are located at the beginning and the end of the LV feeder. Obtained load and PV contribution factors are shown in Table II. PV penetration level for all of the scenarios is 98%. The total installed PV size equals 120.2 kWp, which is the same for all three scenarios. In the next stage, total unmonitored PV generation is required to be estimated using the unsupervised methodology. Using one of the customers' PV output as a solar irradiance proxy located in bus 3, the PV scaling factor is estimated. The gathered data of November 2020 with the resolution of 1 min are utilized to estimate the PV scaling factor, offline. Using the estimated PV scaling factor and RPV output, real-time total PV output can be estimated for December 2020.

2) Real-Time Results: Table III presents the estimation error for one-day period using different methods. According to the results, the proposed PV-ACKF could increase the accuracy of the estimator in areas with PV generation. In scenario 2, most of the PV units are located at the beginning of the feeder, which could decrease the impact of the PV on the conventional ACKF. Since all the PV units are near to reference bus 1, the total network is not affected by the PV, which can improve the accuracy of the estimator. In scenarios 1 and 3, all the buses and branches are affected by the impact of the PV generation; therefore, the results of the ACKF are worse compared with scenario 2. Scenario 3 has the worst results compared with

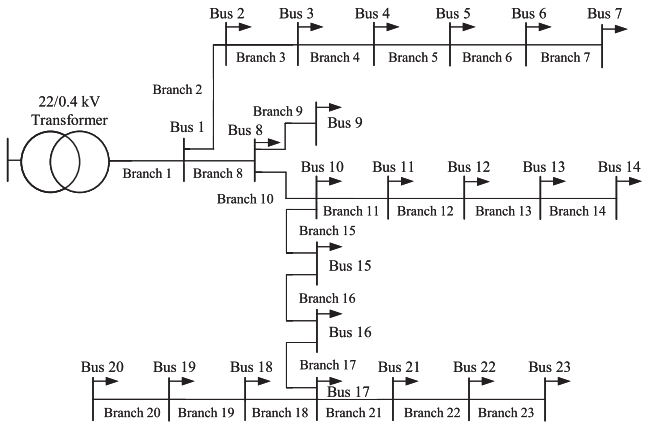

Fig. 5. Australian 23-bus DN [16].

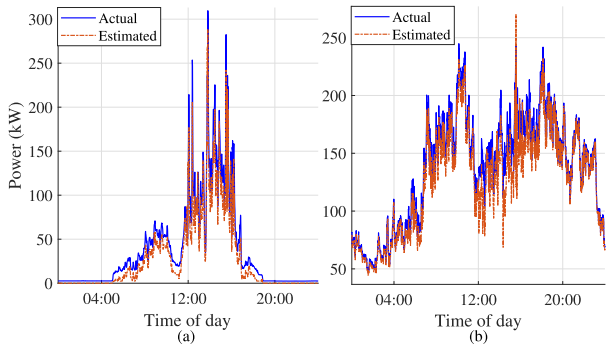

Fig. 6. Comparative results of the estimated total PV and total demand at bus 1 in the 23-bus test system. (a) Total PV. (b) Total demand.

other methods since the majority of the installed PV units are located at the end of the feeder. Using the proposed PV-ACKF, a significant improvement in all the scenarios can be seen. The proposed method is not sensitive to different PV panel locations. The proposed estimator has a superior performance compared with the conventional method in different scenarios.

B. 23-Bus Australian DN

The proposed technique is also applied to an Australian LV network with 23 buses, as illustrated in Fig. 5. Here, 179 loads are connected to the buses. All the information related to the installed PV and load is provided in Table IV. The total installed PV capacity is 543.32 kWp, and the maximum total power consumption on the transformer side is 383.47 kW. The number of measurement points is intentionally kept low in order to reduce observability. Bus 1 voltage and branch 1 current are only available measurements for the state estimator in this network.

1) PV Estimation Results: Using the historical measurement data at bus 1 and one of the customers' PV output (load connected to bus 9) as an irradiance proxy, (10) would calculate the PV scaling factor to estimate the total PV generation on the transformer side. Fig. 6 shows the estimated total demand and PV portions compared with the actual values on the transformer

TABLE III
ESTIMATED VOLTAGE ERROR COMPARING CONVENTIONAL ACKF AND PV-ACKF IN DIFFERENT SCENARIOS IN THE SIX-BUS DN

Bus	Scenario 1				Scenario 2				Scenario 3			
	MAE (%)		MEV (%)		MAE (%)		MEV (%)		MAE (%)		MEV (%)	
	ACKF[12]	**PV-ACKF**	ACKF[12]	**PV-ACKF**	ACKF[12]	**PV-ACKF**	ACKF[12]	**PV-ACKF**	ACKF[12]	**PV-ACKF**	ACKF[12]	**PV-ACKF**
2	0.59	**0.04**	0.43	**0.01**	0.62	**0.04**	0.55	**0.01**	0.59	**0.04**	0.49	**0.01**
3	1.15	**0.10**	1.58	**0.05**	1.10	**0.13**	1.61	**0.06**	1.21	**0.09**	2.00	**0.05**
4	1.56	**0.25**	2.67	**0.13**	1.05	**0.15**	1.69	**0.11**	1.32	**0.20**	2.73	**0.13**
5	1.78	**0.37**	3.43	**0.23**	0.97	**0.17**	1.53	**0.14**	1.40	**0.28**	3.25	**0.19**
6	1.26	**0.12**	1.88	**0.08**	1.01	**0.15**	1.48	**0.10**	1.82	**0.15**	3.85	**0.11**

TABLE IV
23-BUS-TEST-SYSTEM-CONNECTED LOAD AND PV INFORMATION

Bus	Number of loads	PV size (kWp)	Maximum power consumption (kW)	PV Penetration (%)	Bus	Number of loads	PV size (kWp)	Maximum power consumption (kW)	PV Penetration (%)
2	6	16.4	20.48	80.07	13	7	37.4	31.32	119.32
3	8	20.8	32.05	64.897	14	12	44.7	35.63	125.17
4	7	9.6	20.80	46.147	15	7	34.2	28.64	119.37
5	8	8	20.44	39.137	16	7	28.5	28.88	98.39
6	6	11.6	29.37	39.497	17	5	20.1	16.81	119.51
7	8	5.5	17.06	32.227	18	6	22.6	15.32	146.64
8	11	10.4	26.53	39.2	19	8	31.1	19.23	161.03
9	8	15.7	22.06	70.94	20	6	23.6	16.64	140.82
10	8	15.2	22.56	67.30	21	15	73.6	40.16	182.94
11	7	7	17.90	39.09	22	10	56.9	23.12	246.02
12	6	25.9	26.12	98.94	23	13	25.6	35.29	71.97

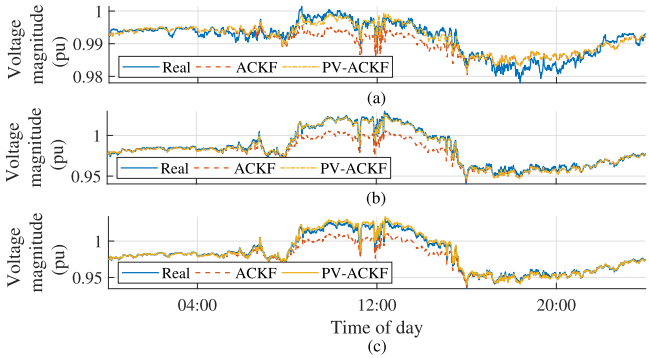

Fig. 7. Real-time results comparing ACKF [12] and PV-ACKF in the 23-bus test system. (a) Bus 6. (b) Bus 13. (c) Bus 22.

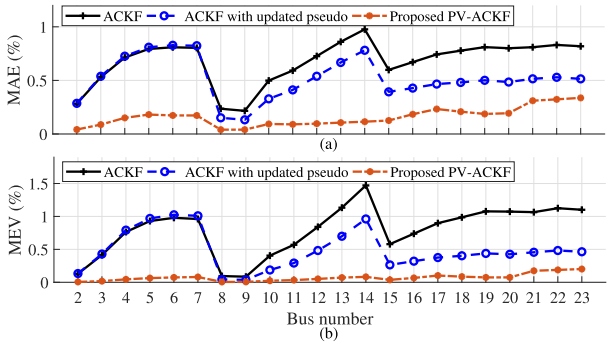

Fig. 8. Voltage magnitude estimation error comparing different methods in the 23-bus test system. (a) MAE (%). (b) MEV (%).

side on December 14, 2020. According to the results, the proposed unsupervised PV estimation technique can estimate PV and demand power with an acceptable error. Using (13) and (15), demand and PV pseudo power measurements are generated to be utilized in the state estimator.

2) Real-Time Results: Fig. 7 shows the performance of the proposed PV-ACKF compared with the conventional ACKF and the actual values for three selected buses on December 21, 2020. According to the results, the conventional ACKF is biased during the mid-day due to the presence of the PV generation, while the proposed PV-ACKF methodology is robust against the generated PV output.

Figs. 8 and 9 illustrate the estimation error for voltage magnitude and angle in different buses over the period of seven days from December 15, 2020 to December 21, 2020. Based on the results, the accuracy of the conventional ACKF is improved if the proposed pseudo generations are used. On the other hand, the proposed PV-ACKF would improve the accuracy of the estimator since the demand portion of the bus current injection is chosen as the system state to fulfill the white noise assumption in the estimator. The results show that bus 14 experiences the

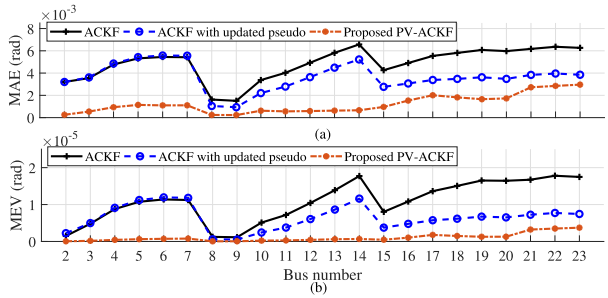

Fig. 9. Voltage angle estimation error comparing different methods in the 23-bus test system. (a) MAE (in radian). (b) MEV (in radian).

highest amount of error. On the other hand, bus 8 has the minimum error compared with other buses. Based on the results, all buses that are far from the reference bus (bus 1) experience higher errors compared with the nearest buses.

Over the entire simulation set, ACKF and PV-ACKF techniques are calculated using identical initial conditions and settings. The average computation time over the period of seven days is about the same for both techniques and equals 7 ms

TABLE V
DIFFERENT PV PENETRATION SCENARIOS

Scenario	1	2	3	4	5	6	7
PV penetration level (%)	70	106	142	177	213	248	284
PV size (kWp)	272.2	408.3	543.3	680.5	816.6	952.7	10888

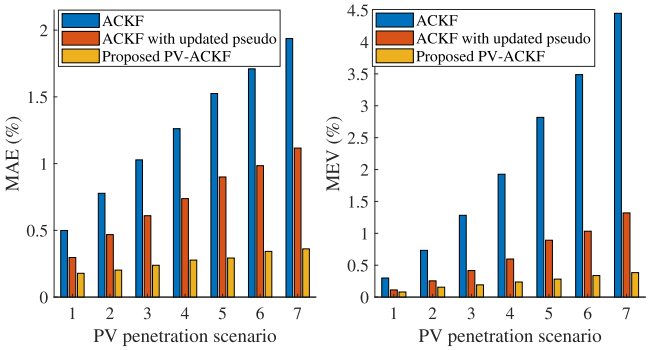

Fig. 10. Voltage estimation error in different PV penetration scenarios. (a) MAE. (b) MEV.

since most of the procedures in the proposed PV-ACKF are implemented offline. Therefore, the PV-ACKF can provide a more accurate estimate with the same computation cost compared with conventional ACKF in areas with high levels of PV penetration.

3) Sensitivity of the PV-ACKF to Different PV Penetration Levels: Here, different PV penetration levels are tested to examine the effectiveness of the proposed PV-ACKF estimator. Seven different PV penetration scenarios are considered to show the impact of PV penetration on the accuracy of the estimator. To generate these PV penetration scenarios, only the installed PV size of each house is changed. PV penetration levels and sizes in different scenarios are presented in Table V. Estimation errors comparing different methods are presented in Fig. 10. According to the results, the ACKF performance is getting worse when the PV penetration level is increasing. The proposed PV-ACKF has a consistent performance with a lower amount of error compared with other methods since the PV separation strategy is utilized in the state and measurement models to reduce the impact of high levels of PV penetration. Based on the results, scenario 1 experiences the lowest amount of error compared with other scenarios since small amounts of PV are generated in this scenario.

C. Modified IEEE-116 LV Test Feeder

In this section, the proposed method is tested on the bigger network and compared with snapshot WLS [21] and recursive EKF [22] estimators to show the effectiveness of the suggested approach. Here, the modified IEEE LV European test feeder with 116 nodes and 55 load buses is used [35], [36]. The topology of the network is shown in Fig. 11, where the voltage at the head of the feeder is set to 416 V. In this case study, the actual data provided by the Energex and Ergon Energy Network are used to create a real case scenario. Here, 165 residential customers are connected to the network, and each load bus consists of

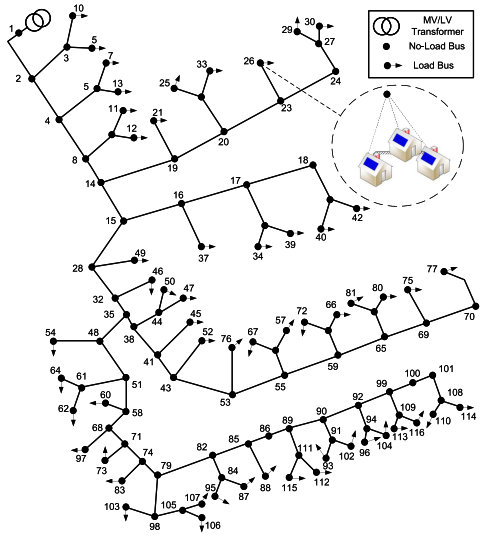

Fig. 11. Modified IEEE-116 LV test feeder.

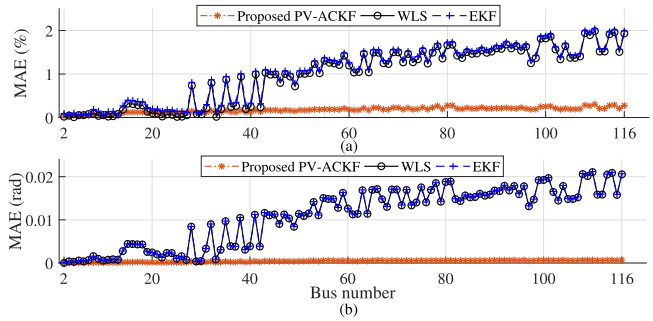

Fig. 12. Voltage MAE comparing different methods in the IEEE-116 LV test feeder. (a) Voltage magnitude. (b) Voltage phase angle.

three individual loads (shown in Fig. 11). All the loads and PV information are provided in Table VI. Like the previous case studies, measurements are only available on the transformer side, and pseudo measurements are generated to run the SE (using customers' billing information and registered PV sizes). One of the customers' PV output connected to bus 105 is randomly chosen as an RPV measurement. The total PV penetration level is 197% in this network.

Here, the SE results are compared with current DSSE methods. We have evaluated the effectiveness of the proposed technique against EKF as a dynamic estimator and WLS as a snapshot estimator. The pseudo data were generated for all these methods using the suggested technique to evaluate all the methods at the same condition. The MAE in bus voltage magnitude and angle across the full simulation set is shown in Fig. 12. According to the results, traditional methods have higher errors compared with the proposed PV-ACKF. Since most of the PV installations are placed down the feeder, the estimation error is larger for those buses at the end of the feeder. On the other hand, the results show the same level of accuracy for both WLS and EKF since both techniques are solving the same set of nonlinear equations representing the link between

TABLE VI
116-BUS-TEST-SYSTEM-CONNECTED LOAD AND PV INFORMATION IN THE LOAD BUSES

Bus	PV size (kWp)	Maximum power (kW)	Bus	PV size (kWp)	Maximum power (kW)	Bus	PV size (kWp)	Maximum power (kW)	Bus	PV size (kWp)	Maximum power (kW)	Bus	PV size (kWp)	Maximum power (kW)
5	0.0	13.19	33	0.0	12.83	52	19.8	16.69	76	27.8	18.55	102	30.6	26.02
7	0.0	14.12	34	0.0	6.05	54	18.3	14.13	77	21.2	19.86	103	32.9	19.66
10	0.0	10.74	37	0.0	6.38	57	25.3	19.39	80	22.1	10.81	104	20.6	19.35
11	0.0	15.84	39	0.0	8.12	60	14.4	18.80	81	27.7	15.25	106	25.2	16.85
12	0.0	10.15	40	12.2	14.87	62	26.7	17.19	83	39.2	29.66	107	27.3	17.39
13	0.0	14.33	42	27.7	21.04	64	11.3	10.81	87	25.2	13.14	110	25.8	16.95
21	0.0	13.22	43	24.3	21.54	66	26.4	14.19	88	17.7	13.41	112	19.1	16.22
25	0.0	14.32	46	13.1	14.26	67	28.2	16.95	93	19.1	16.74	113	32.9	21.10
26	0.0	13.08	47	14.4	10.46	72	29.0	20.42	95	23.5	16.25	114	16.8	12.30
29	0.0	15.90	49	18.0	11.37	73	25.8	20.65	96	25.0	14.15	115	22.5	20.64
30	0.0	7.05	50	26.9	16.22	75	23.7	17.77	97	18.7	15.24	116	22.1	13.84

TABLE VII
COMPUTATIONAL EFFICIENCY AND ACCURACY COMPARISON

Algorithm Type	Maximum Voltage Magnitude MAE (%)	CPU Time (s)		
		Min	Mean	Max
WLS based DSSE	1.9	0.96	1.53	3.92
EKF based DSSE	2.1	0.9	1.35	3.88
Proposed PV-ACKF	**0.3**	**0.65**	**0.93**	**3.12**

the network node voltage and measurements. Using different system covariance matrix, the performance of the dynamic EKF estimator can be worse compared with the traditional snapshot WLS estimator in areas with high levels of PV penetration.

Table VII compares the computational efficiency and the accuracy of the suggested method against conventional WLS- and EKF-based DSSE techniques. The same initial conditions are considered for all of the methods. Furthermore, the maximums of the MAEs are also compared for different methods at all buses. According to the results, the proposed PV-ACKF technique could decrease the maximum voltage error from 2.1% (using the EKF) to 0.3% (using the proposed PV-ACKF). On the other hand, the average computation time is reduced from 1.53 s (using WLS) to 0.93 s (using the proposed PV-ACKF) due to the linear formulation of the PV-ACKF. According to the results, the EKF is better in terms of speed, but it experiences a higher maximum MAE compared with conventional WLS.

V. CONCLUSION

This article presents a new generalized formulation of the ACKF-based DSSE approach that works in poorly monitored LV feeders with high PV penetration levels. Existing unmonitored residential PV installations are violating the accuracy of the voltage estimation techniques in LV feeders by imposing uncertainties and violating noise assumptions. First, the total PV generation is estimated on the transformer side using the differencing strategy. Second, the outputs are used to generate pseudo measurements using the proposed load and PV contribution factors. Finally, the new PV-ACKF estimator is proposed to estimate the bus voltage in the feeder. The proposed technique has been compared with conventional snapshot and dynamic estimators in different scenarios and case studies. According to the results, the suggested approach has a better performance in terms of accuracy and computational efficiency compared with conventional methods.

ACKNOWLEDGMENT

The authors would like to thank Energex and Ergon Energy Network for their help in providing the real dataset to be used in this study.

REFERENCES

[1] F. C. Schweppe and D. B. Rom, "Power system static-state estimation, Part II: Approximate model," *IEEE Trans. Power App. Syst.*, vol. PAS-89, no. 1, pp. 125–130, Jan. 1970.
[2] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*. Boca Raton, FL, USA: CRC Press, 2004.
[3] J. Zhao *et al.*, "Power system dynamic state estimation: Motivations, definitions, methodologies, and future work," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 3188–3198, Jul. 2019.
[4] 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.
[5] D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Hoboken, NJ, USA: Wiley, 2006.
[6] X. Sun, J. Ji, B. Ren, C. Xie, and D. Yan, "Adaptive forgetting factor recursive least square algorithm for online identification of equivalent circuit model parameters of a lithium-ion battery," *Energies*, vol. 12, no. 12, 2019, Art. no. 2242.
[7] H. Madsen, *Time Series Analysis*. London, U.K.: Chapman & Hall/CRC Press, 2007.
[8] A. Louis, G. Ledwich, G. Walker, and Y. Mishra, "Measurement sensitivity and estimation error in distribution system state estimation using augmented complex Kalman filter," *J. Modern Power Syst. Clean Energy*, vol. 8, no. 4, pp. 657–668, 2020.
[9] N. Zhou, D. Meng, Z. Huang, and G. Welch, "Dynamic state estimation of a synchronous machine using PMU data: A comparative study," *IEEE Trans. Smart Grid*, vol. 6, no. 1, pp. 450–460, Jan. 2015.
[10] 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.
[11] J.-H. Teng, "A direct approach for distribution system load flow solutions," *IEEE Trans. Power Del.*, vol. 18, no. 3, pp. 882–887, Jul. 2003.
[12] M. Shafiei, G. Ledwich, G. Nourbakhsh, A. Arefi, and H. Pezeshki, "One-step multilayer unbalanced three-phase DSE for LV distribution networks," *IEEE Syst. J.*, vol. 15, no. 3, pp. 4403–4412, Sep. 2021.
[13] S. Almutairi, Z. Miao, and L. Fan, "Performance of branch-current based distribution system state estimation," in *Proc. IEEE North Amer. Power Symp.*, 2018, pp. 1–6.
[14] H. Liao, "Review on distribution network optimization under uncertainty," *Energies*, vol. 12, no. 17, 2019, Art. no. 3369.
[15] J. Wu, Y. He, and N. Jenkins, "A robust state estimator for medium voltage distribution networks," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1008–1016, May 2013.
[16] M. Shafiei, G. Nourbakhsh, A. Arefi, G. Ledwich, and H. Pezeshki, "Single iteration conditional based DSSE considering spatial and temporal correlation," *Int. J. Elect. Power Energy Syst.*, vol. 107, pp. 644–655, 2019.
[17] M. Shafiei, A. MokariBolhassan, G. Ledwich, G. Nourbakhsh, and A. Arefi, "The impact of loads aggregation and correlation in distribution state estimation," in *Proc. IEEE 10th Int. Workshop Appl. Meas. Power Syst.*, 2019, pp. 1–5.

[18] A. Angioni, T. Schlösser, F. Ponci, and A. Monti, "Impact of pseudo-measurements from new power profiles on state estimation in low-voltage grids," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 1, pp. 70–77, Jan. 2016.

[19] C. Muscas, M. Pau, P. A. Pegoraro, and S. Sulis, "Effects of measurements and pseudomeasurements correlation in distribution system state estimation," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 12, pp. 2813–2823, Dec. 2014.

[20] J. Zhao, L. Mili, and A. Gómez-Expósito, "Constrained robust unscented Kalman filter for generalized dynamic state estimation," *IEEE Trans. Power Syst.*, vol. 34, no. 5, pp. 3637–3646, Sep. 2019.

[21] M. Ayiad, H. Martins, O. Nduka, and B. Pal, "State estimation of low voltage distribution network with integrated customer-owned PV and storage unit," in *Proc. IEEE Milan PowerTech Conf.*, 2019, pp. 1–6.

[22] G. Barchi and D. Macii, "A photovoltaics-aided interlaced extended Kalman filter for distribution systems state estimation," *Sustain. Energy, Grids Netw.*, vol. 26, 2021, Art. no. 100438.

[23] 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.

[24] C. M. Cheung, S. R. Kuppannagari, R. Kannan, and V. K. Prasanna, "Disaggregation of behind-the-meter solar generation in presence of energy storage resources," in *Proc. IEEE Conf. Technol. Sustainability*, 2020, pp. 1–7.

[25] X. Zhang and S. Grijalva, "A data-driven approach for detection and estimation of residential PV installations," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2477–2485, Sep. 2016.

[26] K. Li, F. Wang, Z. Mi, M. Fotuhi-Firuzabad, N. Duić, and T. Wang, "Capacity and output power estimation approach of individual behind-the-meter distributed photovoltaic system for demand response baseline estimation," *Appl. Energy*, vol. 253, 2019, Art. no. 113595.

[27] F. Bu, K. Dehghanpour, Y. Yuan, Z. Wang, and Y. Zhang, "A data-driven game-theoretic approach for behind-the-meter PV generation disaggregation," *IEEE Trans. Power Syst.*, vol. 35, no. 4, pp. 3133–3144, Jul. 2020.

[28] E. Vrettos, E. Kara, E. Stewart, and C. Roberts, "Estimating PV power from aggregate power measurements within the distribution grid," *J. Renewable Sustain. Energy*, vol. 11, no. 2, 2019, Art. no. 023707.

[29] E. C. Kara, C. M. Roberts, M. Tabone, L. Alvarez, D. S. Callaway, and E. M. Stewart, "Disaggregating solar generation from feeder-level measurements," *Sustain. Energy, Grids Netw.*, vol. 13, pp. 112–121, 2018.

[30] W. Li, M. Yi, M. Wang, Y. Wang, D. Shi, and Z. Wang, "Real-time energy disaggregation at substations with behind-the-meter solar generation," *IEEE Trans. Power Syst.*, vol. 36, no. 3, pp. 2023–2034, May 2021.

[31] A. Mokaribolhassan, G. Ledwich, G. Nourbakhsh, A. Arefi, and M. Shafiei, "Solar generation estimation in distribution network using limited measurements," in *Proc. IEEE PES Innov. Smart Grid Technol. Eur.*, 2020, pp. 374–376.

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

[33] M. Shafiei, "Distribution network state estimation, time dependency and fault detection," Ph.D. dissertation, School Elect. Eng. Comput. Sci., Queensland Univ. Technology, St. Lucia, QLD, Australia, 2019.

[34] H. Shaker, H. Zareipour, and D. Wood, "A data-driven approach for estimating the power generation of invisible solar sites," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2466–2476, Sep. 2016.

[35] K. P. Schneider *et al.*, "Analytic considerations and design basis for the IEEE distribution test feeders," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 3181–3188, May 2018.

[36] M. A. Khan and B. Hayes, "A reduced electrically-equivalent model of the IEEE european low voltage test feeder." 2020. [Online]. Available: https://dx.doi.org/10.21227/0d2n-j565