Local Hybrid Linear State Estimation for Electric Power Systems Using Stream Processing

Kang Sun, Student Member, IEEE, Manyun Huang, Member, IEEE, Zhinong Wei, Member, IEEE, Yuzhang Lin, Member, IEEE, and Guoqiang Sun, Member, IEEE

Abstract—The increasing penetration of renewable energy resources with highly fluctuating outputs has placed increasing concern on the accuracy and timeliness of electric power system state estimation (SE). Meanwhile, we note that only a fraction of system states fluctuate at the millisecond level and require to be updated. As such, refreshing only those states with significant variation would enhance the computational efficiency of SE and make the fast-continuous update of states possible. However, this is difficult to achieve with conventional SE methods, which generally refresh the states of the entire system every 4-5 s. In this context, we propose a local hybrid linear SE framework using stream processing, in which the synchronized measurements received from phasor measurement units (PMUs), and trigger/timing-mode measurements received from remote terminal units (RTUs) are used to update the associated local states. Moreover, the measurement update process efficiency and timeliness are enhanced by proposing a trigger measurementbased fast dynamic partitioning algorithm for determining the areas of the system with states requiring recalculation. In particular, non-iterative hybrid linear formulations with both RTUs and PMUs are employed to solve the local SE problem. The timeliness, accuracy, and computational efficiency of the proposed method are demonstrated by extensive simulations based on IEEE 118-, 300-, and 2383-bus systems.

Index Terms—Local hybrid linear state estimation, stream processing, phasor measurement units, trigger/timing-mode measurements, fast dynamic partitioning, timeliness.

I. INTRODUCTION

THE secure and economic operation of electric power systems relies heavily on the accurate and timely state information provided by state estimation (SE) [1]. The raw data employed for SE derive mainly from remote terminal units (RTUs) [2]. Recently, the deployment of phasor measurement units (PMUs) in power systems is continually increasing [3], [4], which can provide complex voltage and line currents in a synchronized manner, resulting in enhanced monitoring and control capabilities above those of RTUs [5]– [7]. As a result, a range of new PMU-based applications can be developed, such as event detection, system frequency response, oscillation analysis, and state estimation [8], [9]. However, PMUs cannot yet guarantee the observability of the entire electric power system [10], [11]. Therefore, RTUs can be expected to play a leading role in power system SE for the foreseeable future. In general, RTUs provide both timing and trigger measurement modes. In the timing mode, the scan interval of RTUs is approximately 5 s, and all available measurements are uploaded each time. In the trigger mode, the measured values of RTUs are uploaded immediately once the measured data changes beyond a predefined triggering threshold. By contrast, PMU only provide timing measurement mode with the sampling rate of approximately 30 or 60 frames/s. Thanks to the high sampling rate of PMUs, we can simulate the trigger measurement mode and recover missing data through subsequent data processing programs [12], [13]. Under hybrid RTU/PMU measurement conditions, measured data can be divided into the following two categories.

- Timing measurements (TiMs): Timing measurements, which are obtained periodically, include RTU measurements in the timing mode and corresponding PMU measurements. Moreover, the time-skew problem caused by different sampling rate of RTUs and PMUs can be dealt with data buffers [14], [15].
- Trigger measurements (TrMs): Trigger measurements, which are generated and uploaded in real time, include RTU measurements in the trigger mode and PMU measurements in the aforementioned simulated trigger mode.

In this context, only TiMs are used by conventional SE. However, such an estimator can only provide estimated result as the same rate as the RTUs, such as once every 5 s. At the same time, we note that the position of the TrMs reflect the areas with significant variation. Hence, if SE is carried out only within these local areas in parallel, the computational cost will be greatly reduced, and SE can be executed at a high frequency due to the real-time nature of TrMs. In this regard, stream processing has been demonstrated to be an efficient approach for analyzing the large scale multiple data streams generated by measurement devices [16]–[18]. This approach are able to organizes distributed computing resources and expected to promote the use of real-time information for improving the tracking performance of SE [19].

Various methods have been proposed to develop efficient SE algorithms. For example, semidefinite programming [20], [21] reformulates the SE problem using rank relaxation, while other approaches use bilinear formulations [22], [23]. However, these algorithms do not utilize PMUs and require a number of iterations to converge. Non-iterative SE algorithms can be developed utilizing PMUs alone [24], [25] if the electric power

This work was supported by the National Key Research and Development Program of China under Grant 2018YFB0904500.

K. Sun, M. Huang, Z. Wei and G. Sun are with the College of Energy and Electrical Engineering, Hohai University, Nanjing 210098, China (email: sunkang.real@qq.com; hmy_hhu@yeah.net; wzn_nj@263.net; hhusunguoqiang@163.com).

Y. Lin is with the Department of Department of Electrical and Computer Engineering, University of Massachusetts Lowell, Lowell MA 01854, USA (e-mail: yuzhang_lin@uml.edu). DOI: 10.17775/CSEEJPES.2020.06000

system is observable by PMUs with sufficient redundancy [26]. However, a more realistic solution is using hybrid SE that utilizes both RTU and PMU measurements [27]. In this context, RTU and PMU measurements can be processed jointly [2], [27], or separately [7], [28]–[30]. Moreover, a various methods for addressing bad data in hybrid SE have been proposed [31]– [33]. Recently, a non-iterative linear SE using RTUs alone has been demonstrated to realize ultra-fast SE using a complex form of calculation [34], and its extension [35] introduce a real form SE using hybrid RTU/PMU measurements.

In addition to the above discussed issues, existing SE methods generally refresh the states of the entire electric power system which lead to the high computational cost [36]. Efforts have been made to address this issue to some extent by the development of multi-area SE (MASE), which divides the network into small sub-areas and conducts SE for all sub-areas in parallel [36]–[38]. For example, one method applied partitions according to the refresh rate of RTUs and PMUs in the network and the SE execution time [38]. Other partitioning approaches are data driven [36], [39] and have employed factor graphs [37]. However, the subareas in the above MASE approaches are fixed prior, which severely limits their flexibility. Moreover, SE must be applied to all of the sub-areas, resulting in high computational burden.

The above-discussed issues are addressed in the present study by proposing a non-iterative local hybrid linear SE (LHLSE) framework based on stream processing. This framework, which is composed of a fast dynamic partitioning (FDP) algorithm and a hybrid linear SE algorithm (HLSE), is illustrated in Fig. 1. Here, TiMs are directly utilized by HLSE algorithm to provide estimated values of the entire power system peridically, while TrMs are exploited by the complete LHLSE algorithm to update the local area states in real time. The mean processing time of the LHLSE algorithm t must be considered in the measurement processing steps. As shown in Fig. 1, the measurements are grouped according to t, which will be processed by SE together. The stream data pipe acts as a buffer, and is used to cache measurements data that cannot be processed in time. Unlike the conventional SE framework, which only updates system states periodically, the proposed stream processing framework can provide estimated results at any time when the measured data are uploaded. The main contributions of this work are as follows.

- We propose a generalized hybrid linear SE algorithm, denoted herein as the HLSE algorithm, which can conduct conventional periodic global SE and real-time aperiodic local SE.
- 2) We propose a fast dynamic partitioning (FDP) algorithm for aperiodic SE, which can dynamic adjust the areas with states requiring recalculation according to the TrMs in real time, and then instantly obtain the local SE model of the related subareas.
- 3)We propose an LHLSE algorithm that can quickly perform multiple local SE calculations between two global HLSE processes in parallel. With stream processing and TrMs, the system states are updated at a high frequency, improving the tracking performance of SE to the millisecond level.



Fig. 1. Stream processing framework for the LHLSE algorithm.

The remainder of this paper is structured as follows. Section II describes the proposed HLSE algorithm in detail. Section III introduces the FDP algorithm and the LHLSE implementation. Section IV presents and analyzes simulation results obtained using the proposed method for three electric power test systems. Section V concludes the paper.

II. HYBRID LINEAR STATE ESTIMATION

This section briefly presents an existing RTU-based linear state estimation (LSE) algorithm [34] and then formulates the HLSE algorithm.

A. LSE Process

The relationship between system states and measurements can be represented as a linear system of equations by defining a pseudo complex current measurement $(.)^{local}$ as

$$\dot{I}_{ft}^{local} = I_{ft} e^{j\theta_{ft}}.$$
(1)

Here, I denotes the current measurement, subscript ft denotes a transmission line from bus f to bus t, $\theta_{ft} = tan^{-1}(-Q_{ft}/P_{ft})$, P and Q denote the active and reactive power measurements, respectively. With reference to the slack bus, the complex current measurement can be expressed as

$$\dot{I}_{ft} = \dot{I}_{ft}^{local} e^{j\delta_f},\tag{2}$$

where δ_f is the unknown phase angle of the complex voltage at bus f with reference to the slack bus. The relationships between complex state variables with (pseudo) complex measurements are described by rewriting the measurement function as follows [34].

1) Voltage measurements: Complex voltages are given as

$$U_f = U_f e^{j\delta_f},\tag{3}$$

where U_f and δ_f are state variables and U_f is related to the measured voltage amplitude at bus f as

$$U_f^{meas} = U_f + \varepsilon_{U_f}.$$
 (4)

2) Current flow measurements: Current flow measurements are given as

$$\dot{I}_{ft} = \dot{I}_{ft}^{local} e^{j\delta_f} = (\dot{y}_{ft0} + \dot{y}_{ft}) \dot{U}_f + (-\dot{y}_{ft}) \dot{U}_t, \quad (5)$$

where \dot{y}_{ft0} and \dot{y}_{ft} are the shunt admittance and series admittance of transmission line *f*-*t* at bus *f*, respectively.



Fig. 2. Simple 3-bus system with voltage and power measurements [40].

3) Injected current measurements: Injected Current Measurements are given as

$$\dot{I}_f = \dot{I}_f^{local} e^{j\delta_f} = \sum_{t \in L_f} \dot{I}_{ft}.$$
(6)

Here, \dot{I}_{ft} is the complex current vector of line f-t, \dot{Y} is the admittance matrix, and \dot{U} is the complex voltage vector.

Combining both voltage and current measurements according to the above formalizations yields a linear system of equations that can be summarized in a compact form as

$$H_A x + \varepsilon_{A_1} = z_A,\tag{7}$$

where H_A is a matrix of size $m \times (2n-1)$ containing measurements of all buses except the slack bus whose measurements are in z_A , $x = [\dot{U}_1, \dots, \dot{U}_n, e^{j\delta^2}, \dots, e^{j\delta^n}]$ is a vector of state variables, and ε_A is a vector of complex measurement errors.

The structures of (7) are illustrated by considering the simple 3-bus system shown in Fig. 2, where bus 1 is specified as the slack bus. This yields the following H_A , z_A and x.

$$H_{A} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -U_{2} & 0 \\ 0 & 0 & 1 & 0 & -U_{3} \\ \hline \dot{y}_{120} + \dot{y}_{12} & -\dot{y}_{12} & 0 & 0 & 0 \\ 0 & \dot{y}_{230} + \dot{y}_{23} & -\dot{y}_{23} & -\dot{I}_{23}^{local} & 0 \\ 0 & -\dot{y}_{32} & \dot{y}_{320} + \dot{y}_{32} & 0 & -\dot{I}_{32}^{local} \\ \hline -\dot{y}_{31} & 0 & \dot{y}_{310} + \dot{y}_{31} & 0 & -\dot{I}_{31}^{local} \\ \hline Y_{21} & Y_{22} & Y_{23} & -\dot{I}_{2,inj}^{local} & 0 \\ Y_{31} & Y_{32} & Y_{33} & 0 & -\dot{I}_{3,inj}^{local} \end{bmatrix},$$

$$z_{A} = \begin{bmatrix} U_{1} & 0 & 0 & \dot{I}_{12}^{local} & 0 & 0 & 0 & \dot{I}_{1}^{local} & 0 & 0 \end{bmatrix}^{T}, \quad (9)$$

$$x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T = \begin{bmatrix} \dot{U}_1 & \dot{U}_2 & \dot{U}_3 & e^{j\delta_2} & e^{j\delta_3} \end{bmatrix}^T, (10)$$

where x_1 denotes the complex voltage vector and x_2 denotes the voltage phase-angle vector. Then, H_A and z_A can be treated as a whole by merging z_A into H_A , yielding

$$H_{zA} = \begin{bmatrix} n & n \\ \check{H}_{11} & \check{H}_{z12} \\ \check{H}_{21} & \check{H}_{z22} \\ \check{H}_{31} & \check{H}_{z32} \end{bmatrix} m_U m_U m_{I_{fl}} = \begin{bmatrix} E_A & -U \\ \dot{Y}^{local} & -\dot{I}^{local}_{ft} \\ \dot{Y} & -\dot{I}^{local}_{fc} \end{bmatrix} .$$
(11)

Here, \check{H}_{11} , \check{H}_{21} , and \check{H}_{31} are all constant matrices, \check{H}_{z12} , \check{H}_{z22} , and \check{H}_{z32} contain all the measurements, and E_A is

part of a unit matrix that is of a size $n \times n$ according to the voltage measurement configuration. Here, the following two-stage scheme [34] is employed to ensure the linearity of the SE.

Stage 1: Obtain complex measurements

The estimated state variables can be obtained via

$$\hat{x} = (H_A^T H_A)^{-1} H_A^T z_A.$$
 (12)

Then, the complex measurements can be obtained as

$$m = -\begin{bmatrix} \check{H}_{z12}^T & \check{H}_{z22}^T & \check{H}_{z12}^T \end{bmatrix}^T e^{j\delta}.$$
 (13)

Stage 2: Execute complex SE

The model is now formulated as a linear WLS estimation by incorporating all the voltage and current measurements into the measurement vector m as

$$Jx_1 + \varepsilon_{A_2} = m, \tag{14}$$

where $J = \begin{bmatrix} H_{11}^T & H_{12}^T & H_{13}^T \end{bmatrix}^T$, $m = \begin{bmatrix} \dot{U}_f^T & \dot{I}_{ft}^T & \dot{I}_f^T \end{bmatrix}^T$, and $x_1 = \begin{bmatrix} \dot{U}_1 \cdots \dot{U}_n \end{bmatrix}^T$ denotes the complex voltage vector. The linear WLS estimation is given by

$$\hat{x}_1 = \left[J^T R^{-1} J\right]^{-1} J^T R^{-1} m,$$
(15)

where $R = E(\varepsilon \cdot \varepsilon^T)$ is a diagonal covariance matrix of measurements.

B. HLSE Process

The complex bus voltage and line current measurements of PMUs can be linearly related to system state variables as follows.

1) Voltage measurements: We relate complex voltage measurements to state variables as

$$\dot{U}_f^{meas} = \dot{U}_f + \varepsilon_{\dot{U}_f}.$$
(16)

2) Current flow measurements: Complex current flow measurements are related to state variables as

$$\dot{I}_{ft}^{meas} = (\dot{y}_{ft0} + \dot{y}_{ft}) \dot{U}_f + (-\dot{y}_{ft}) \dot{U}_t + \varepsilon_{\dot{I}_{ft}}.$$
 (17)

According to the above PMU measurement formulations, we obtain a linear matrix equation, that is

$$H_B x + \varepsilon_B = z_B, \tag{18}$$

which can be combined with (14) into the following single equation.

$$Hx + \varepsilon = z, \tag{19}$$

where the matrix H and vector z can be rendered as

$$H = \begin{bmatrix} \check{H}_{11}^T & \check{H}_{21}^T & \check{H}_{31}^T \end{bmatrix} \check{H}_{41}^T & \check{H}_{51}^T \end{bmatrix}^T,$$
(20)

$$z = \left[\begin{array}{c} m^T & z_B^T \end{array} \right]^T. \tag{21}$$

Here, H_{41} and H_{51} are structurally similar to H_{11} and H_{21} . Then the SE process can be completed in a linear and noniterative manner utilizing both RTUs and PMUs.



Fig. 3. Example 12-bus system with one U-area.

III. LOCAL STATE ESTIMATION BASED ON DYNAMIC PARTITIONING

The proposed LHLSE algorithm is able to determine the areas whose states requiring recalculation according to the TrMs, and then perform local HLSE for each area. As shown in Fig. 3, one original area (O-area) and several update areas (U-areas) can be formed, and the buses in U-areas are divided into two categories:

1) Internal bus: An internal bus is one associated to updated TrMs, and the collection of internal buses forms several internal areas (I-area).

2) Boundary bus: A boundary bus belongs to the U-area but not to the I-area, it connected to the internal buses by a line which do not have any TrM update.

The cluster of internal bus and boundary bus forms the Uarea, while all buses except internal bus represent the O-area. Among them, only the states of U-areas must be recalculated. Since the boundary buses do not have any measurements update, it is only used for merging phase angles, and only the states of I-areas are updated in the SE process. It is worthwhile to note that any two U-areas are always disconnected from each other in the FDP algorithm, such that the calculations for each area are independent.

The proposed methodology is illustrated by the 12-bus system shown in Fig. 3 with two U-areas. At a particular instant, the TrMs of bus 2, 4, 10, and branch 1-4 are updated. Then, buses 1, 2, 4, and 10 are denoted as internal buses, which form the I-areas, and buses 3, 5, 9, and 11 represent the boundary buses. Obviously, the updated measurements are too few to satisfy the requirements of the whole system observability and to achieve an acceptable noise filtering effect. Hence, the results of the preceding SE process must be utilized as the pseudo-measurements of the current SE process, which are considered to be sufficiently accurate under the assumption that the TrM and the actual value change nearly synchronously. Thus, the measurement redundancy required for SE can be ensured in the U-area. The above partition logic is implemented by FDP algorithm in an efficient way and discussed in Section III-B.

A. U-Area State Estimator Formulation

The SE process is completed in each U-area independently and in parallel with all other U-areas. After obtaining the Uarea states, the voltage phase angles of the U-area must be merged with those of the O-area. The phase angle difference

Algorithm 1: Internal Bus Merge Algorithm
Input: internal bus matrix V
Output: I-area matrix \mathcal{V}
1 initialize $i = 1$;
2 repeat
3 if $V_i^R \neq 0$ (the <i>i</i> th bus is an internal bus) then
4 extend two layers outward from the <i>i</i> th internal
bus: $e_1 = \Psi\{N_{V^R}^R\}$, and $e_2 = \Psi\{N_{e_1}^R\} - V_i^R$;
5 yielding extended internal buses $e_{in} = \Psi\{V_{e_2}^R\}$;
6 if $e_{in} \neq 0$ (extended internal buses exist) then
7 merge the internal buses: $V_i^R = V_i^R e_{in};$
8 delete merged internal buses: $V_{e_{in}}^R = 0;$
9 else
10 move to the next bus: $i = i + 1$;
11 end
12 else
13 move to the next bus: $i = i + 1$;
14 end
15 until $i = n$;
16 Delete the zero rows of the matrix $V, V \leftarrow V$;

between the slack bus of the U-area and of the O-area is obtained as

$$\Delta \delta = \frac{1}{n_b} \sum_{b} \delta_b^U - \delta_b^O, \qquad (22)$$

where the superscripts U and O denote the U-area and O-area, respectively, and subscript b indexes the n_b boundary buses in the designated U-area. As such, we can merge the phase angles of the I-area to those of the O-area by the following formulation.

$$\delta_I^O = \delta_I^U - \Delta \delta, \tag{23}$$

where δ_I^U denotes the phase angle in the I-area with reference to the slack bus of the U-area, and δ_I^O denotes the phase angle in the I-area with reference to the slack bus of the O-area.

B. FDP Algorithm

In this subsection, the U-area determination approach and the linear model partition approach are presented. All the U-areas are obtained according to the measurement upload position, and HLSE models are established within each Uarea through model partitioning, without recalculating the coefficient matrix. In addition, the computations of the FDP algorithm are almost bit-operations, which are ultra-efficient in the implementation. By contrast, the conventional method is not able to make real-time decisions on which states need to be updated, nor can it obtain local SE results as fast as the proposed method. Moreover, the Jacobian matrix of the conventional SE model must be recalculated in each iteration, which will cause unacceptable computational costs in realtime applications, even if part of the state variables is fixed in advance.

1) Determine the U-areas: Let N denote a Boolean network matrix composed of elements with values of 0 or 1 that has the same distribution of non-zero elements as admittance



Fig. 4. Boolean matrixes of the FDP algorithm for the 12-bus system when the measurements of bus 2, 4, 10, and branch 1-4 are updated.

matrix Y, and M denote a Boolean measurement update status matrix, where each element of M corresponds to the measurement update status of a bus or branch of the network, and an element with a value of 1 indicates a measurement update for that object; otherwise, the value is 0. We note that only a few measurements are updated at any given time. Therefore, M is relatively sparse. In addition, let $\Psi\{(\cdot)\}$ denote the bitwise OR operator applied between each row of the matrix corresponding to its argument, $(\cdot)|(\cdot)$ denote the bitwise OR operator applied to two matrices with the same dimension, $(\cdot)_v^R$ denote a vector that includes all the entries in the vth row of the matrix corresponding to its argument, and v is a Boolean index vector, where element values of 1 represent those corresponding elements of a given vector that are retained; otherwise, the elements of the vector are discarded. For example, vectors $A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \end{bmatrix}$ and $v = [1 \ 0 \ 1 \ 1 \ 0]$ produce a final vector $A_v^R = [1 \ 3 \ 4]$. The above definitions are now applied in the following three steps of the FDP algorithm.

Step 1: Determine the internal bus index vector v_{in} according to the measurement update status matrix through

$$v_{in} = \Psi\{ M \} | \Psi\{ M^T \}.$$
 (24)

Step 2: Determine the I-areas. Let $V = diag(v_{in})$ denote the internal bus matrix, where $diag(\cdot)$ denotes the diagonalization operation. Then, I-area matrix V is obtained in the internal bus merge algorithm (Algorithm 1), which will be illustrated with a 12-bus system later. Each row of V represents all the internal buses in the respective U-area.

Step 3: Determine the U-area matrix as follows.

$$\mathcal{T}_i^R = \Psi\{N_{\mathcal{V}_i^R}^R\}.$$
(25)

Here, \mathcal{T} has a size $n_U \times n$, where n_U is equal to the number of U-areas. Each row of \mathcal{T} represents an U-area, such that \mathcal{T} has the same size as \mathcal{V} .

Note that the maximum time complexity of the above Iarea merge algorithm is $\mathcal{O}(2n)$. The computations in all the three steps are bit-operations, which are very efficient for most computer languages, and the elements of all matrices and vectors employed in the FDP algorithm are Boolean terms, which require very little storage space. For the 12-bus system in Fig. 3, we can establish the M matrix as shown in Fig. 4(a) when the measurements of bus 2, 4, 10 and branch 1–4 are updated. The green and gray areas represent the location of element 1 of network matrix N. According to (24), internal bus index vector v_{in} can be obtained as

As shown in Fig. 4(c) and (27), two U-areas are established. The first U-area consists of bus 1, 2, 3, 4, and 5, among which buses 1, 2, and 4 are internal buses, the second U-area consists of bus 9, 10, and 11, among which bus 10 is internal bus.

2) Partition the hybrid linear model: Once the U-area and I-area are obtained, the global model can be directly partitioned into sub-area models. There is no need to recalculate each coefficient matrix of HLSE algorithm within each U-area, which can greatly improve the efficiency of the state estimator. According to (8), (11), (14), and (18), H_A , J, H_B , and z_A are sub-matrixes of H_{zA} when the estimated results of the preceding SE are accepted as pseudo-measurements. Therefore, we can consider the condition defined by H_{zA} alone, and expand it into the full form given by (7) and (19).

Let H_{zA}^U denote the elements of H_{zA} in (12) that pertain to the U-areas. This matrix is obtained by conducting a partial extraction according to a row index vector r_A and a column index vector c. Here, we analyze only the first U-area because



Fig. 5. Impact of PMUs proportion on estimation error.



Fig. 6. Impact of PMUs proportion on computation time.

all of the U-areas operate independently and in parallel. The internal bus vectors and all bus vectors of this first U-area are represented herein as $v = V_1^R$ and $t = \mathcal{T}_1^R$, respectively. Each network branch measurement is incident to two buses. Therefore, the corresponding network graph is an undirected graph, and a bus-branch measurement incident matrix B can be defined as

$$B_{ij} = \begin{cases} 1 & \text{if branch measurement } j \text{ is incident to bus } i \\ 0 & \text{otherwise} \end{cases}$$

It should be noted that B is composed of constants that require no recalculation when the network topology is determined. The vectors r_A and c can be obtained as

$$r_A = \begin{bmatrix} t & \Psi\{B_v^R\} & v \end{bmatrix}, c &= \begin{bmatrix} t & t \end{bmatrix},$$
(28)

where the three elements of r_A correspond to voltage measurement, current flow measurement and injected current measurement, respectively, and the two elements of c correspond to complex voltage x_1 and phase angle x_2 , respectively. By conducting a partial extraction according to r_A and c, we have

$$H_{zA}^U = H_{zA}(r_A, c).$$
 (29)

Accordingly, the linear model in (7) and (19) can be solved locally.



(b) Phase angle estimate error

Fig. 7. Comparison of HLSE estimation error for the different systems.

IV. CASE STUDIES

The effectiveness, accuracy, and computational efficiency of the proposed method were evaluated by its application to 1000 Monte Carlo simulations of IEEE 118-, 300-, and 2383bus systems [41]. Independent zero mean Gaussian errors were added to all the measurements. The standard deviations of the RTU measurements were 0.005 p.u. for voltage measurements and 0.01 p.u. for power measurements. The standard deviations of the PMU measurements was 0.001 p.u. All simulations were conducted using a PC with an Intel Core i7-8750H CPU operating at 2.21 GHz with 16 GB of RAM. The performances of both the HLSE and LHLSE algorithms were evaluated in comparison with the performances obtained by the conventional WLS algorithm [40], [42]-[44] and real formulation linear SE (RFLSE) [35]. The proportion of PMU devices in the system are artificially specified in all simulations, but it should be noted that the locations of PMU devices are random in each measurement set. The primary performance metric employed was the mean absolute error (MAE), which is defined as

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |x_i^{esti} - x_i^{true}|,$$
 (30)

where x_i^{esti} and x_i^{true} are the estimated and true value of state variables (i.e., the voltage magnitude and phase angle), respectively. In addition, the computation times of the various SE algorithms were compared.

A. Performance of the HLSE Algorithm

1) Impact of PMUs proportion: The SE accuracy of the proposed HLSE algorithm for the IEEE 118-bus system under different proportions of PMUs is presented in Fig. 5. Here, the result of 1000 Monte Carlo simulations are given by blue box plots and the mean value is given by red line. As expected, the accuracy and stability of SE increases with an increasing proportion of PMUs. The corresponding computation time

TABLE I
MAE VALUES OF SE ALGORITHMS FOR THE IEEE 118-BUS SYSTEM

Algorithm	Voltage Magnitude (p.u.)	Phase Angle (rad)
WLS [40]	$5.51 imes 10^{-4}$	5.97×10^{-4}
RFLSE [35]	1.28×10^{-4}	1.14×10^{-4}
HLSE	1.18×10^{-4}	$1.08 imes 10^{-4}$

 TABLE II

 Computation Time of SE Algorithms for Different Systems (s)

System	WLS [40]	RFLSE [35]	HLSE	
118-bus	0.0171	0.0029	0.0021	
300-bus	0.0489	0.0067	0.0041	
2383-bus	0.2664	0.0843	0.0419	

 TABLE III

 Partitioning Results for the IEEE 118-Bus System

U-area	Size	Internal Bus	Boundary Bus
1	13	1, 2, 3, 11, 12, 14	4, 5, 7, 13, 15, 16, 117
2	51	34, 35, 37, 39, 43-49, 69, 77, 78, 80-83, 92-98, 100	19, 33, 36, 38, 40, 42, 50, 51, 54, 66, 68, 70, 75, 76, 79, 84, 85, 89, 91, 99, 101-104, 106

of the HLSE algorithm is plotted in Fig. 6. These results demonstrate that the impact of the additional PMUs is almost negligible. The SE accuracy of the proposed HLSE algorithm with a PMU proportion of 50% is plotted in Fig. 7 in terms of the number of simulations obtaining a given MAE value for the three different IEEE power systems.

2) Comparison with conventional SE: The average MAE values obtained by the three SE algorithms considered over 1000 simulation cases are listed in Table I for the IEEE 118bus system, where the proportion of PMUs was 50% for the HLSE and real form linear SE (RFLSE) algorithms [35]. It is evident that the HLSE algorithm outperforms the WLS and RFLSE algorithms in terms of SE accuracy. In addition, the computation times required by the three SE algorithms for the three systems are listed in Table II. Here, the HLSE algorithm demonstrably outperforms the WLS and RFLSE algorithm in terms of computational efficiency.

B. Performance of the LHLSE Algorithm

The effectiveness of the proposed LHLSE approach is validated through comparative simulation settings. It is assumed that in the previous instant, TiMs are uploaded which make the entire system observable; and in the current instant,TrMs are uploaded which make only a portion of the system observable. Results of two simulation cases are compared:

1) Performing HLSE for the entire system using TiMs from the previous instant;

2) Performing HLSE for the entire system using TiMs from the previous instant; then performing LHLSE (FDP + HLSE) to update the local areas using TrMs from the current instant.



Fig. 8. Visualization of network matrix N for the IEEE 118-bus system.



(b) Phase angle estimation error in the overall system.

Fig. 9. Comparisons of estimation errors for the IEEE 118-bus system.

Here, the active load and reactive load of each selected bus was increased by 10%. In addition, the measurement data in each simulation case were generated by dynamic power flow [45], which can handle power change and does not depend on the selected slack bus, to provide more realistic simulation results. In practice, once TrMs are updated, several U-area and I-area will be formed by FDP algorithm quickly.

1) Effectiveness of the LHLSE algorithm: We first compare the performances of the conventional SE framework and the LHLSE algorithm with a PMU proportion of 0% for the IEEE 118-bus system with the stated load fluctuations at buses 2, 45, 46, 94, 95, and 96 of the system. The partitioning result of the IEEE 118-bus system by the FDP algorithm is listed in Table III and visualized in Fig. 8. Here, U-areas are given by orange circles and I-areas are given by red circles on the basis of the network matrix given by blue circles. The results in the figure indicate that only a small proportion of states in the overall system requires recalculation and updating. The SE errors obtained by the above two simulation cases within the entire system are given in Fig. 9. We can conclude that the FDP algorithm accurately identifies the bus positions at which state variables change, and the SE process in the U-

 TABLE IV

 Normalized Residual in Bad Data Identification for the IEEE

 118-Bus System

Iteration	I_{33-37}	P_{101}	\dot{I}_{81-80}	Other Measurements
1	38.809	10.013	87.452	28.769
2	36.827	8.7442	-	9.9878
3	-	8.6244	-	5.9363
4	-	-	-	2.9318

TABLE V MAE VALUES OF SE ALGORITHMS IN I-AREA FOR DIFFERENT SYSTEMS

System	Algorithm	Voltage Magnitude (p.u.)	Phase Angle (rad)
118-bus	Conventional	1.32×10^{-4}	8.53×10^{-3}
	LHLSE	9.76×10^{-5}	5.48×10^{-5}
300-bus	Conventional	1.43×10^{-3}	3.57×10^{-1}
	LHLSE	7.52×10^{-5}	5.71×10^{-5}
2383-bus	Conventional	1.08×10^{-4}	1.09×10^{-2}
	LHLSE	4.88×10^{-5}	3.93×10^{-5}

area yields estimates with greatly reduced MAE values than those obtained by the conventional SE framework.

2) Bad data processing: The ability to detect erroneous measurements, as well as further identify and suppress them is a significant function of a state estimator [40]. In this regard, Chi-squares test is utilized for bad data detection. If any bad data is detected, further identification and elimination processing will be carried out using Largest Normalized Residual (LNR) test [40]. In order to evaluate the performance of bad data processing, various bad data simulations have been carried out for the IEEE 118-bus system successfully. As an example, when the PMU proportion is 50% and there are load fluctuations at buses 2, 45, 46, 94, 95, and 96, the RTU current measurement for line 33-37, the RTU injected active power measurement for bus 101, and the PMU phasor current measurement for line 81-80 are set to -100%, 50%, and 0% of the raw data, respectively, to simulate erroneous measurements. As expected, bad data are detected successfully. When the identification threshold is chosen as 3 [40], the LNR test process is presented in Table IV. The normalized residual of the three erroneous measurements are listed in the first three column, and the last column provides the largest normalized residual of the remaining normal measurements for comparison purposes. It can be seen that after three iterations, all the bad data have been correctly identified and eliminated.

3) Accuracy and efficiency of the LHLSE aogorithm: We next compare the performances of the conventional SE framework and the LHLSE algorithm with a PMUs proportion of 50% for all three IEEE power systems with the stated load fluctuations applied at 5% of the system buses selected randomly. The average MAE values of the estimated state variables obtained for the three systems are listed in Tables V. In addition, we compare the computation times required by the LHLSE and HLSE algorithms for the three IEEE systems in

 $\label{eq:TABLE VI} \begin{array}{c} TABLE \mbox{ VI} \\ COMPUTATION TIME OF SE ALGORITHMS FOR DIFFERENT SYSTEMS (S) \end{array}$

System	HLSE	LHLSE(FDP + HLSE)
118-bus 300-bus 2383-bus	$\begin{array}{c} 2.17 \times 10^{-3} \\ 4.15 \times 10^{-3} \\ 4.18 \times 10^{-2} \end{array}$	$\begin{array}{l} 9.51(4.43+5.08)\times 10^{-4}\\ 3.01(0.97+2.04)\times 10^{-3}\\ 9.29(4.91+4.38)\times 10^{-3} \end{array}$

Table VI. Note that the SE process in each U-area is completed in parallel. Hence, only the maximum computation time of all areas is listed. It should be noted that the PMU device locations differ for Table II and Table VI lead to slightly different computation times. These results demonstrate that the proposed LHLSE algorithm is computationally efficient and accurate, particularly for the large-scale system.

V. CONCLUSION

This paper developed an LHLSE framework with stream processing that enabled the use of PMU measurements and trigger/timing-mode RTU measurements. The timeliness and computational efficiency of measurement update was enhanced by proposing a new FDP algorithm that divided the system into several U-areas whenever measurement values were updated, and the designed HLSE algorithm utilizing both RTU and PMU measurements was applied to each U-area in parallel. The stream processing framework of the LHLSE algorithm guarantees the real-time tracking and correction of information for the entire power system states in a few milliseconds. The timeliness, accuracy, and computational efficiency of the proposed method were demonstrated by extensive simulations based on IEEE 118-, 300-, and 2383-bus power systems.

REFERENCES

- J. Zhao, G. Zhang, K. Das, G. N. Korres, N. M. Manousakis, A. K. Sinha, and Z. He, "Power system real-time monitoring by using PMUbased robust state estimation method," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 300–309, Jan. 2016.
- [2] M. Göl, A. Abur, and F. Galvan, "Rapid tracking of bus voltages using synchro-phasor assisted state estimator," in *IEEE PES ISGT Europe* 2013, Oct. 2013, pp. 1–5.
- [3] S. Xu, H. Liu, T. Bi, and K. E. Martin, "A high-accuracy phasor estimation algorithm for PMU calibration and its hardware implementation," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 1–1, Jul. 2020.
- [4] M. Huang, Z. Wei, J. Zhao, R. A. Jabr, M. Pau, and G. Sun, "Robust ensemble kalman filter for medium-voltage distribution system state estimation," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 7, pp. 4114–4124, Jul. 2020.
- [5] J. Li, S. Xu, H. Liu, and T. Bi, "A high-accuracy and low-complexity phasor estimation method for pmu calibration," *CSEE Journal of Power* and Energy Systems, pp. 1–9, Jul. 2020.
- [6] X. Zhao, H. Zhou, D. Shi, H. Zhao, C. Jing, and C. Jones, "On-line PMU-based transmission line parameter identification," *CSEE Journal* of Power and Energy Systems, vol. 1, no. 2, pp. 68–74, Jun. 2015.
- [7] N. M. Manousakis, G. N. Korres, J. N. Aliprantis, G. P. Vavourakis, and G.-C. J. Makrinas, "A two-stage state estimator for power systems with PMU and SCADA measurements," in 2013 IEEE Grenoble Conference, Jun. 2013, pp. 1–6.
- [8] J. De La Ree, V. Centeno, J. S. Thorp, and A. G. Phadke, "Synchronized phasor measurement applications in power systems," *IEEE Transactions* on Smart Grid, vol. 1, no. 1, pp. 20–27, Jun. 2010.
- [9] J. H. Chow, Online monitoring of dynamics with PMU data. Wiley, Aug. 2016.

- [10] Y. Zhang and J. Wang, "Towards highly efficient state estimation with nonlinear measurements in distribution systems," *IEEE Transactions on Power Systems*, pp. 2471–2474, May 2020.
- [11] J. Zhao, J. Qi, Z. Huang, A. P. S. Meliopoulos, A. Gomez-Exposito, M. Netto, L. Mili, A. Abur, V. Terzija, I. Kamwa, B. Pal, and A. K. Singh, "Power system dynamic state estimation: Motivations, definitions, methodologies, and future work," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 3188–3198, Jul. 2019.
- [12] P. Gao, M. Wang, S. G. Ghiocel, J. H. Chow, B. Fardanesh, and G. Stefopoulos, "Missing data recovery by exploiting low-dimensionality in power system synchrophasor measurements," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1006–1013, Mar. 2016.
- [13] D. Osipov and J. H. Chow, "PMU missing data recovery using tensor decomposition," *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4554–4563, Nov. 2020.
- [14] V. Murugesan, Y. Chakhchoukh, V. Vittal, G. T. Heydt, N. Logic, and S. Sturgill, "PMU data buffering for power system state estimators," *IEEE Power and Energy Technology Systems Journal*, vol. 2, no. 3, pp. 94–102, Sep. 2015.
- [15] Q. Zhang, Y. Chakhchoukh, V. Vittal, G. T. Heydt, N. Logic, and S. Sturgill, "Impact of PMU measurement buffer length on state estimation and its optimization," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1657–1665, May 2013.
- [16] W. X, Z. Y, and L. H, "Reliable stream data processing for elastic distributed stream processing systems," *Cluster Computing*, 2019.
- [17] M. H. Javed, X. Lu, and D. K. Panda, "Cutting the tail: Designing high performance message brokers to reduce tail latencies in stream processing," in *IEEE International Conference on Cluster Computing* (*CLUSTER*). IEEE, Sep. 2018, Conference Proceedings, p. 223–233.
- [18] J. Tao, M. Umair, M. Ali, and J. Zhou, "The impact of internet of things supported by emerging 5G in power systems: A review," *CSEE Journal* of Power and Energy Systems, vol. 6, no. 2, pp. 344–352, Jun. 2020.
- [19] M. Dias De Assunção, A. Da Silva Veith, and R. Buyya, "Distributed data stream processing and edge computing: A survey on resource elasticity and future directions," *Journal of Network and Computer Applications*, vol. 103, pp. 1–17, Feb. 2018.
- [20] Y. Lan, H. Zhu, and X. Guan, "Fast nonconvex SDP solvers for largescale power system state estimation," *IEEE Transactions on Power Systems*, vol. 35, no. 3, pp. 2412–2421, May 2020.
- [21] H. Zhu and G. B. Giannakis, "Power system nonlinear state estimation using distributed semidefinite programming," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 6, pp. 1039–1050, Dec. 2014.
- [22] A. Gomez-Exposito, C. Gomez-Quiles, and A. D. L. V. Jaen, "Bilinear power system state estimation," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 493–501, Feb. 2012.
- [23] W. Zheng, W. Wu, A. Gomez-Exposito, B. Zhang, and Y. Guo, "Distributed robust bilinear state estimation for power systems with nonlinear measurements," *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 499–509, Jan. 2017.
- [24] M. Göl and A. Abur, "A fast decoupled state estimator for systems measured by PMUs," *IEEE Transactions on Power Systems*, vol. 30, no. 5, pp. 2766–2771, Sep. 2015.
- [25] J. Lin, Y. Su, Y. Cheng, C. Lu, L. Zhu, H. Huang, and Y. Liu, "A robust complex-domain state estimator using synchrophasor measurements," *International Journal of Electrical Power & Energy Systems*, vol. 96, no. MAR., pp. 412–421, Mar. 2018.
- [26] M. Göl and A. Abur, "A hybrid state estimator for systems with limited number of PMUs," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1511–1517, May 2015.
- [27] G. Valverde, S. Chakrabarti, E. Kyriakides, and V. Terzija, "A constrained formulation for hybrid state estimation," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1102–1109, Aug. 2011.
- [28] M. Zhou, V. A. Centeno, J. S. Thorp, and A. G. Phadke, "An alternative for including phasor measurements in state estimators," *IEEE Transactions on Power Systems*, vol. 21, no. 4, pp. 1930–1937, Nov. 2006.
- [29] A. Simoes Costa, A. Albuquerque, and D. Bez, "An estimation fusion method for including phasor measurements into power system real-time modeling," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1910–1920, May 2013.
- [30] N. M. Manousakis and G. N. Korres, "A hybrid power system state estimator using synchronized and unsynchronized sensors," *International Transactions on Electrical Energy Systems*, vol. 28, no. 8, pp. 1–19, Aug. 2018.
- [31] M. Y. Ccahuana, F. Schmidt, and M. C. De Almeida, "Analysis of bad data detection in power system state estimators considering PMUs," in 2015 IEEE Power Energy Society General Meeting, Jul. 2015, pp. 1–5.

- [32] J. Chen and A. Abur, "Placement of PMUs to enable bad data detection in state estimation," *IEEE Transactions on Power Systems*, vol. 21, no. 4, pp. 1608–1615, Nov. 2006.
- [33] G. N. Korres and N. M. Manousakis, "State estimation and bad data processing for systems including PMU and SCADA measurements," *Electric Power Systems Research*, vol. 81, no. 7, pp. 1514–1524, Jul. 2011.
- [34] A. S. Dobakhshari, S. Azizi, M. Paolone, and V. Terzija, "Ultra fast linear state estimation utilizing SCADA measurements," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 2622–2631, Jul. 2019.
- [35] A. S. Dobakhshari, M. Abdolmaleki, V. Terzija, and S. Azizi, "Robust hybrid linear state estimator utilizing SCADA and PMU measurements," *IEEE Transactions on Power Systems*, pp. 1–1, Aug. 2020.
- [36] B. Zargar, A. Angioni, F. Ponci, and A. Monti, "Multiarea parallel data-driven three-phase distribution system state estimation using synchrophasor measurements," *IEEE Transactions on Instrumentation and Measurement*, no. 9, p. 6186–6202, Sep. 2020.
- [37] Y. Guo, L. Tong, W. Wu, H. Sun, and B. Zhang, "Hierarchical multi-area state estimation via sensitivity function exchanges," *IEEE Transactions* on *Power Systems*, vol. 32, no. 1, pp. 442–453, Jan. 2017.
- [38] A. Sharma, S. C. Srivastava, and S. Chakrabarti, "Multi area state estimation for smart grid application utilizing all SCADA and PMU measurements," 2014 IEEE Innovative Smart Grid Technologies - Asia, ISGT ASIA 2014, pp. 525–530, May 2014.
- [39] D. Zhang, X. Han, and C. Deng, "Review on the research and practice of deep learning and reinforcement learning in smart grids," *CSEE Journal* of Power and Energy Systems, vol. 4, no. 3, pp. 362–370, Sep. 2018.
- [40] A. Abur and A. Gomez-Exposito, Power System State Estimation: Theory and Implementation. Portland: Ringgold Inc, 2004, vol. 28.
- [41] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "MAT-POWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, Feb. 2011.
- [42] F. Schweppe, "Power system static-state estimation, part III: Implementation," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-89, no. 1, pp. 130–135, Jan. 1970.
- [43] F. Schweppe and D. Rom, "Power system static-state estimation, part II: Approximate model," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-89, no. 1, pp. 125–130, Jan. 1970.
- [44] F. Schweppe and J. Wildes, "Power system static-state estimation, part I: Exact model," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-89, no. 1, pp. 120–125, Jan. 1970.
- [45] R. Ramanathan, H. Ramchandani, and S. A. Sackett, "Dynamic load flow technique for power system simulators," *IEEE Transactions on Power Systems*, vol. 1, no. 3, pp. 25–30, Aug. 1986.



Kang Sun (S'21) received the B.S. degree from the College of Energy and Electrical Engineering, Hohai University, Nanjing, China, in 2019, where he is currently pursuing the Ph.D. degree.

His current research interests include power system state estimation, cyber-physical systems, and high-performance computing.



Manyun Huang (S'15, M'19) received the B.S. and Ph.D. degree from the College of Energy and Electrical Engineering, Hohai University, Nanjing, China, in 2014 and 2019, respectively.

From September 2017 to September 2018, she was a joint Ph.D. Student with RWTH Aachen University, Aachen, Germany. She is currently an Associate Professor with the College of Energy and Electrical Engineering, Hohai University. Her current research interests include the theory and algorithms of power system state estimation, robust

state estimation, and Kalman filter.



Zhinong Wei (M'14) received the B.S. degree from the Hefei University of Technology, Hefei, China, in 1984, the M.S. degree from Southeast University, Nanjing, China, in 1987, and the Ph.D. degree from Hohai University, Nanjing, in 2004.

He is currently a Professor of electrical engineering with the College of Energy and Electrical Engineering, Hohai University. His current research interests include power system state estimation, integrated energy systems, smart distribution systems, optimization and planning, load forecasting, and

integration of distributed generation into electric power systems.



Yuzhang Lin (S'13, M'18)received the B.S. and M.S. degrees from Tsinghua University, Beijing, China, in 2012 and 2014, respectively, and the Ph.D. degree from Northeastern University, Boston, MA, USA, in2018.

He is currently an Assistant Professor with the Department of Electrical Computer Engineering, University of Massachusetts, Lowell, MA,USA. His current research interests include modeling, monitoring, cyber–physical security, and data analysis of smart grids.



Guoqiang Sun (M'14) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Hohai University, Nanjing, China, in 2001, 2005, and 2010, respectively.

He was a Visiting Scholar with North Carolina State University, Raleigh, NC, USA, from 2015 to 2016. He is currently an Associate Professor with the College of Energy and Electrical Engineering, Hohai University. His current research interests include power system analysis and economic dispatch and optimal control of integrated energy systems.