Centralized Identification of Imbalances in Power Networks with Synchrophasor Data

Tirza Routtenberg Member, IEEE and Yonina C. Eldar, Fellow, IEEE

Abstract—The problem of bus imbalance identification in a three-phase power network using a phasor measurement unit (PMU) is considered. We propose new algorithms to identify and localize imbalances occurring at any location in the power grid, based on single-phase PMU data. First, we develop a technique using the minimum description length (MDL) criterion that is carried out at the control center. The centralized MDL methodology is time consuming and has computational complexity that grows exponentially with the network size. Therefore, we next develop a Projected Orthogonal Matching Pursuit (POMP) algorithm, which is a low-complexity method for bus imbalance identification in large-scale power networks implemented at the control center. POMP is a computationally efficient compressive sensing technique that exploits the sparse structure of the voltage measurements. The proposed methods are validated through three case studies: a two-port π-model, an IEEE-14 bus system, and an IEEE 118-bus system. Simulations show that our networked algorithms, MDL and POMP, obtain improved performance over local bus-level identification techniques. The performance of POMP is close to that of centralized MDL, with the advantage of being applicable to large-scale networks.

Index Terms—Signal detection, phasor measurement unit (PMU), unbalanced power system, smart grid, multi-area state estimator, networked identification of imbalances, minimum description length (MDL), Orthogonal Matching Pursuit (OMP).

I. INTRODUCTION

Three-phase power systems frequently suffer from imbalances, such as load imbalances [1]. These may be precursors to more serious contingencies leading to blackouts, excessive losses, insulation degradation, and production interruptions [2], [3]. In addition, system imbalance degrades fault location accuracy [4] and state estimation performance, both in terms of bias [5] and error covariance matrices [6]. Thus, imbalance must be detected and compensated for. The use of modern devices, such as phasor measurement units (PMUs), is highly desirable for modern power systems as it facilitates rapid detection of contingencies and faults and provides accurate state and frequency estimation [1].

Traditionally, most energy management system (EMS) functions, such as state estimation, are carried out at the central unit, based on the full network data. While three-phase measurements of power and voltages are available at the local measurement device level, and, in principle, can be transmitted, only the positive sequence components are typically reported to the control center, due to communication and processing limitations. Existing methods for voltage imbalance detection are based on two- or three-phase sequences and are performed locally at the bus level. The derivation of enhanced imbalance identification methods that use single-phase networked data is crucial for obtaining an accurate system model and high power quality. Such processing methods should be able to absorb the increasing number of available measurements, while retaining low communication capacity and low computational complexity.

The symmetrical components transformation converts three phase-measurements into a positive, a negative and a zero sequence. In balanced scenarios, it produces a significant value only in the positive sequence, since all phases are rotated to be added in-phase [2], [7]. Thus, monitoring the power system and developing state estimation methods are usually based on the positive sequence and assume that the network is balanced and can be represented by a single-phase equivalent (e.g. [1] and [8]–[10]). Recently, new state and frequency estimation approaches have also been proposed for different unbalanced system models and scenarios [11]–[16].

Classical detection of imbalances is based on measures such as the voltage unbalance factor (VUF) [17], [18], the phase voltage unbalance factor [19], and the complex VUF [20], [21]. A smart meter for voltage imbalance detection is described in [22]. In [23] a hazard-based model was introduced to predict the three-phase imbalance based on historical status data. New bus-level parametric imbalance detection methods that outperform VUF in terms of probability of error have recently been proposed, based on the generalized likelihood ratio test (GLRT) [13], [24] and the generalized locally most powerful test (GLMP) [25], [26]. For networked identification of imbalances, a non-parametric VUF approach is derived in [3], based on distribution system state estimation (DSSE) in a network. A GLRT-based test and cumulative sum (CUSUM) approach have been developed in [27] for change-detection of the imbalance condition. The VUF, GLRT, GLMP, and networked-DSSE methods are all based on two- or three-phase data that are usually not available at the control center in current networks. In contrast, our proposed parametric techniques employ networked single-phase data, and thus the need for additional communication and processing costs is avoided.

In this paper, we formulate the identification of unbalanced buses as a model order selection problem. It is well known that compressive sensing (CS) can be used to treat model order selection and parameter estimation problems with significant reduction in computational complexity [28]. Reconstruction of sparse signals and CS techniques have become very popular in the last two decades (see e.g. [29]–[31] and references therein). In power systems, CS algorithms [32] have been proposed in the context of multi-line outage identification in transmission networks [33]–[35], gross error identification [36], bad data

T. Routtenberg is with Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel. Email: tirzar@bgu.ac.il. Y. C. Eldar is with the Department of Electrical Engineering, Technion-Israel Institute of Technology, Haifa 32000, Israel. Email: yonina@ee.technion.ac.il. The research of T. Routtenberg was partially supported by the ISRAEL SCIENCE FOUNDATION (ISF), Grant No. 1173/16.
denotes the Moore-Penrose pseudo-inverse given by \( A^\dagger = (A^H A)^{-1} A^H \). The operators Real\{\cdot\} and Imag\{\cdot\} denote the real and imaginary parts of their arguments, respectively. Given a vector \( \mathbf{u} = [u_1, \ldots, u_n]^T \) and a set \( S = \{i_1, \ldots, i_k\} \subset \{1, \ldots, n\} \) of integers, \( \mathbf{u}_S = [u_{i_1}, \ldots, u_{i_k}]^T \) denotes the subvector of \( \mathbf{u} \) indexed by \( S \). Similarly, given a matrix \( \mathbf{Q} \) and a set \( S \), \( \mathbf{Q}_S \) is the submatrix whose columns are indexed by \( S \). The Euclidean \( \ell_2 \)-norm is denoted by \( \| \cdot \|_2 \).

Finally, variables are cataloged in the nomenclature Table.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a power network that is observable through PMUs installed at all buses or at optimal locations [44]. The proposed method aims to identify the unbalanced buses among the measured buses, based only on the positive sequence at the control center. For instance, in the IEEE 14-bus test system, which is shown in Fig. 1 as a single line diagram, the PMUs are located at buses 2, 6, 7, 9, which results in an observable system [44]. Similar to [2], loading imbalances are found at the buses with load 4, 5 and 9. The goal is to use the PMU voltage and current measurements to locate the unbalanced buses.

In this section, we present the mathematical model of the PMU positive sequence output. In Subsection II-A, we describe the conventional model (e.g., [1] and [45]) for single-bus measurements. In Section II-B, we derive a model for the total measurements at the control center, obtained from the \( M \) PMU-equipped buses, based on the single-bus measurement models.
A. Single-bus PMU measurement model

The voltages in a three-phase power system at each bus are assumed to be pure sinusoidal signals of frequency $\omega_0 + \Delta$, where $\omega_0$ is the known nominal frequency (100π or 120π) and $\Delta$ is the frequency deviation. Due to system inertia, the frequency can be assumed identical over the network of measurement nodes. The magnitudes and phases of the three voltages of the $n$th bus are denoted by $V_{a,m}, V_{b,m}, V_{c,m}$ $\geq 0$ and $\varphi_{a,m}, \varphi_{b,m}, \varphi_{c,m} \in [0, 2\pi)$, respectively, for any $m = 1, \ldots, M$. Therefore, the sampled three-phase voltages can be represented in discrete time form as

$$
v_{a,m}[n] = V_{a,m} \cos\left(\frac{\omega_0 + \Delta}{\omega_0} n + \varphi_{a,m}\right),
$$

(1)

$$
v_{b,m}[n] = V_{b,m} \cos\left(\frac{\omega_0 + \Delta}{\omega_0} n + \varphi_{b,m}\right),
$$

(2)

$$
v_{c,m}[n] = V_{c,m} \cos\left(\frac{\omega_0 + \Delta}{\omega_0} n + \varphi_{c,m}\right),
$$

(3)

for any $n = 0, \ldots, N - 1$, where $\gamma \triangleq \frac{\Delta}{\omega_0}$ and the sampling rate is $N$ times per cycle of the nominal frequency, $\omega_0$. For the sake of simplicity, the sequences are presented here without noise. The noise statistic is discussed in the networked model.

In current systems, while all three-phase voltages and currents are monitored by PMUs, usually only the positive sequence is transmitted to the control center [2], [8]. The positive sequence voltage, aka the “space vector” [46], is calculated from three-phase voltages by using the symmetrical component transformation (11) pp. 63-67, [47]:

$$
v_{+,[n]}[n] = \frac{1}{3} \left( v_{+,[n]}[n] + \alpha v_{b,[n]}[n] + \alpha^2 v_{c,[n]}[n] \right),
$$

(4)

for any $n = 0, 1, \ldots, N - 1$, where $\alpha = e^{j2\pi/3}$. By substituting (1)-(3) in (4) and using the trigonometrical identity $\cos\alpha = \frac{1}{2}e^{j\pi/3} + \frac{1}{2}e^{-j\pi/3}$, we obtain the voltage signal at the $n$th bus:

$$
v_{+,[n]}[n] = e^{j\gamma \omega_0 n} C_{+,[n]} + e^{-j\gamma \omega_0 n} C_{-,[n]},
$$

(5)

for any $n = 0, \ldots, N - 1$ and any bus $m = 1, \ldots, M$, where

$$
C_{+,m} \triangleq \frac{1}{6} \left( V_{a,m} e^{j\varphi_{+,m}} + \alpha V_{b,m} e^{j\varphi_{-,m}} + \alpha^2 V_{c,m} e^{j\varphi_{-,m}} \right),
$$

(6)

$$
C_{-,m} \triangleq \frac{1}{6} \left( V_{a,m} e^{j\varphi_{+,m}} + \alpha V_{b,m} e^{j\varphi_{+,m}} + \alpha^2 V_{c,m} e^{j\varphi_{+,m}} \right)
$$

(7)

are the positive and negative sequence phasors at the $n$th bus.

The $n$th bus is perfectly balanced or symmetrical if $V_{a,m} = V_{b,m} = V_{c,m}$ and $\varphi_{a,m} = \varphi_{b,m} + \frac{2\pi}{3} = \varphi_{c,m}$. By substituting these values in (7), it can be verified that for a perfectly balanced bus $C_{-,m} = 0$. Therefore, the model for a balanced system is given by

$$
v_{+,m}[n] = e^{j\gamma \omega_0 n} C_{+,m}, \quad n = 0, \ldots, N - 1.
$$

(8)

For a balanced system, the positive sequence, $v_{+,m}[n]$, contains the same information as (1)-(3). Consequently, only the positive sequence components are typically reported to the control center. However, in the presence of imbalances and, faults, the negative and zero sequences [1] are nonzero [8] and the three-phase measurements from (1)-(3) cannot be reconstructed from the positive sequence.

Existing methods for imbalance detection are based on two- or three-phase sequences. That is, in addition to the positive sequence in (4), the negative sequence

$$
v_{-,m}[n] = \frac{1}{3} \left( v_{a,m}[n] + \alpha^2 v_{b,m}[n] + \alpha v_{c,m}[n] \right),
$$

(9)

$n = 0, \ldots, N - 1$, is required for implementing local and networked methods (e.g. [3], [13], [17], [18], [24]– [26]). The different models in (5) and (8) indicate that voltage imbalance is detectable in the time domain, even locally, by using only the positive sequence model with at least $N = 2$ measurements. In this work we develop local and networked methods that are based on single-phase data. To the best of our knowledge, this is the first published model that shows that unbalanced buses can be identified by using only the single-phase positive sequence.

B. Networked measurement model

The measurement set at the control center consists of the positive-sequence measurements obtained by the $K$ PMUs. In addition to measuring the $K$ bus voltages where the PMUs are installed, PMUs measure the $L$ current phasors in lines connected to these buses. Thus, the noisy measurement vector that is received at the control center is given by [8]

$$
z[n] = Bv_{+,n} + w[n], \quad n = 0, \ldots, N - 1,
$$

(10)

where $z[n] \in C^{(K+L)}$ contains $K$ voltage and $L$ current measurements, and the system matrix is given by $B \triangleq \begin{bmatrix} \mathbf{A} & \mathbf{0} \end{bmatrix} \in C^{(K+L)\times M}$, in which $\mathbf{A}$ is a $K \times M$ matrix with columns from an identity matrix representing the places with a PMU and $\mathbf{Y} \in C^{L\times M}$ is a function of admittance and topology factors (e.g. Chapter 7 in [1]). The voltage vector, $v_{+,n} = [v_{+,1}[n], \ldots, v_{+,M}[n]]^T$, is the system state vector at time $n$, in which $v_{+,m}[n]$ is the voltage at the $m$th bus at time $n$ for any $m = 1, \ldots, M$. The noise vector sequence $w[n]$, $n = 0, \ldots, N - 1$, is assumed to be a time-independent zero-mean complex circularly symmetric Gaussian noise sequence with known covariance matrix, $\sigma^2 \mathbf{I}_{K+L}$. In the following, we assume that the system is perfectly observable by these PMUs, i.e. $\mathbf{B}$ is a full column rank matrix and $K + L \geq M$.

By substituting $v_{+,m}[n]$ from (5) in (10), one obtains

$$
z[n] = e^{j\gamma \omega_0 n} \mathbf{B}_c + e^{-j\gamma \omega_0 n} \mathbf{B}_c + w[n],
$$

(11)

where $\mathbf{c}_+ \triangleq [C_{+,1}, \ldots, C_{+,M}]^T$ and $\mathbf{c}_- \triangleq [C_{-,1}, \ldots, C_{-,M}]^T$ include the positive and conjugate negative sequence phasors, respectively, as defined in (6) and (7). In particular, for a perfectly balanced system, $\mathbf{c}_- = 0$ and, thus, the model in (11) is reduced to

$$
z[n] = e^{j\gamma \omega_0 n} \mathbf{B}_c + w[n], \quad n = 0, \ldots, N - 1.
$$

(12)

If some of the buses are unbalanced, then only the elements of the vector $\mathbf{c}_-$ that are related to the balanced buses are equal to zero. In this case, the measurement model in (11) can be reformulated accordingly, as shown in the next section. We consider the following problem: given the observation vector, $z[n], n = 0, \ldots, N - 1$, determine the imbalance condition of each of the $M$ buses in the network. Thus, there are in total $P = 2^M$ different possible events, which correspond to all the
possible subsets of unbalanced buses from the $M$ buses in the network.

III. NETWORKED IDENTIFICATION VIA THE MDL APPROACH

In this section, we formulate our problem as a model selection problem with multiple hypotheses, where the hypotheses correspond to the number and location of unbalanced buses. Model selection refers to finding the most suitable choice of the model from a set of candidates that provides the “best” description of the observations.

A. The hypothesis-testing problem

The task of identifying the subset of unbalanced buses is equivalent to finding the support of the indices of nonzero elements of $c_-$. This can be done by selecting from amongst the multiple hypotheses:

$$
\mathcal{H}_i : z[n] = e^{j\frac{\pi}{2}\omega_0 n}Bc_+ + e^{-j\frac{\pi}{2}\omega_0 n}B(i)c_-(i) + w[n],
$$

(13)

for $i = 0, 1, \ldots, P$ and $n = 0, \ldots, N-1$, where $B(i)$ is the measurement model given by (12).

The nested composite multiple hypothesis-testing problem in (13) can be approximated by using the asymptotically-consistent ML model selection criterion [39], [40], which is widely employed in signal and array processing. The MDL method chooses the hypothesis $\mathcal{H}_i$ which minimizes the sum of two terms: the likelihood term for data encoding, evaluated at the ML points, and a penalty function that inhibits the number of free parameters of the model from becoming very large. For our model, the MDL term is

$$
\tau_{\text{MDL}}^{(i)} = -L \left( \hat{c}_+^{(i)}, \hat{c}_-^{(i)} \right) + \frac{n^{(i)}}{2} \log N, \quad i = 1, \ldots, P,
$$

(14)

where $\hat{c}_+^{(i)}$ and $\hat{c}_-^{(i)}$ are the ML estimates of $c_+$ and $c_-$, respectively, under the $i$th hypothesis. Based on (13), the measurement log-likelihood function under each hypothesis, $i$, is given by:

$$
L \left( c_+, c_-^{(i)} \right) = \text{const} - \frac{1}{\sigma^2} \times \sum_{n=0}^{N-1} \left\| z[n] - e^{j\frac{\pi}{2}\omega_0 n}Bc_+ + e^{-j\frac{\pi}{2}\omega_0 n}B(i)c_-^{(i)} \right\|^2_2
$$

(15)

i = 1, \ldots, P, where “const” is a constant term which is independent of the unknown vectors, $c_+$ and $c_-^{(i)}$, and of the hypothesis $\mathcal{H}_i$. The number of free (real) unknown parameters, $n^{(i)}$, is equal to the dimension of the complex vectors $c_+$ and $c_-$. That is, under the $i$th hypothesis

$$
n^{(i)} = 2M + 2|S_i|,
$$

(16)

where $|S_i|$ is the number of unbalanced buses under hypothesis $i$, i.e. the length of $|c_-|_{S_i}$.

B. Networked ML state estimation

In order to implement the ML rule from (14), we first develop the ML estimators under each hypothesis.

By equating the derivative of (15) with respect to $c_+$ to zero, one obtains (e.g. Chapter 4 in [48] and [49])

$$
\hat{c}_+^{(i)} = B^\dagger \left( z_+ - \eta B(i)\hat{c}_-^{(i)} \right),
$$

(17)

where

$$
z_+ = \frac{1}{N} \sum_{n=0}^{N-1} e^{-j\frac{\pi}{2}\omega_0 n}z[n], \quad \eta = \frac{1}{N} \sum_{n=0}^{N-1} e^{-j2\frac{\pi}{2}\omega_0 n}.
$$

Similarly, by equating the derivative of (15) with respect to $c_-^{(i)}$ to zero, we have

$$
\hat{c}_-^{(i)} = \left( B(i) \right)^\dagger \left( z_- - \eta^* B\hat{c}_+^{(i)} \right),
$$

(18)

where

$$
z_- = \frac{1}{N} \sum_{n=0}^{N-1} e^{j\frac{\pi}{2}\omega_0 n}z[n].
$$

In particular, for the perfectly balanced hypothesis $H_0$, as described in (12), the ML estimator from (17) and (18) are reduced to $\hat{c}_+^{(i=0)} = B^\dagger z_+$ and $\hat{c}_-^{(i=0)} = 0$, which coincides with existing results on the ML estimator for balanced systems (e.g. [1]).

Substituting (17) in (18) results in

$$
\hat{c}_-^{(i)} = \left( B(i) \right)^\dagger \left( z_- - \eta^* B\hat{c}_+^{(i)} \right).
$$

(19)

Therefore, assuming $\eta \neq 1$ and using the fact that $(B(i)\dagger)^\dagger B\dagger B(i) = I_{|S_i|}$, (19) implies that

$$
\hat{c}_-^{(i)} = \frac{1}{1 - |\eta|^2} (B(i)\dagger)^\dagger \left( z_- - \eta^* B\hat{c}_+^{(i)} \right).
$$

(20)

Plugging (20) into (17) and using $B(i)\dagger B(i)\dagger B(i) = B(i)(B(i)\dagger)^\dagger$, results in

$$
\hat{c}_+^{(i)} = \frac{1}{1 - |\eta|^2} B \left( G_i z_+ - \eta B\hat{c}_+^{(i)} (B(i)\dagger)^\dagger z_- \right),
$$

(21)

where $G_i = (1 - |\eta|^2)I + |\eta|^2 B(i)(B(i)\dagger)^\dagger$.

The estimators in (20) and (21) are functions of the frequency deviation, $\Delta$. If $\Delta$ is unknown, then by equating the derivative of (15) with respect to $\Delta$ to zero, we can obtain its ML estimator. For real-time applications, this estimate, which is based on searching the parameter space, can be replaced by low-complexity frequency estimation methods [1], [11], [51]. In particular, in this paper we use the estimate [1]:

$$
\Delta = \frac{\omega_0}{NM\gamma} \sum_{m=1}^{M} \sum_{n=0}^{N-2} \left( \omega_{n+m}[n+1] \right) - \left( \omega_{n+m}[n] \right),
$$

(22)

where $\omega(\cdot)$ denotes the angle of its argument. Then, the state estimators in (20) and (21) are updated accordingly. In this case, i.e. unknown frequency, the number of free parameters from (16), $n_i$, must be incremented by one.

\footnote{Typical frequency-deviation values in power systems satisfy $|\Delta| \leq 0.1\pi$ [50]. Thus, it can be verified that $|\eta| \ll 1$ for $N > 1$.}
C. The MDL method

Substituting the ML estimator from (17) in (15), we obtain

\[
L \left( \hat{c}_+^i, \hat{c}_-^i \right) = -\frac{1}{\sigma^2} \sum_{n=0}^{N-1} \left| \left| z[n] - e^{j\gamma} \frac{\omega_0}{\omega} \mathbf{B} \mathbf{B}^\dagger z_+^i \right| \right|^2 \\
+ \frac{N(1 - |\eta|^2)}{\sigma^2} \left| \left| (\mathbf{B}^i) \hat{c}_+^i \right| \right|^2,
\]

where we used the fact that \((\mathbf{B}^i)^H \mathbf{B} \mathbf{B}^\dagger \mathbf{B}^i = (\mathbf{B}^i)^H \mathbf{B}^i\).

Removing the terms that are independent of the hypothesis and substituting (16) and (23) into (14), the MDL criterion for selection of the set of unbalanced buses is based on choosing the hypothesis which maximizes

\[
T_i^{(\text{MDL})} = \frac{N(1 - |\eta|^2)}{\sigma^2} \left| \left| (\mathbf{B}^i) \hat{c}_+^i \right| \right|^2 - |S_i| \log N,
\]

for all \(i = 1, \ldots, P\), where the last equality is obtained by substituting (20) and using \((\mathbf{B}^i)^H (\mathbf{B}^i)^{\dagger} \mathbf{B} \mathbf{B}^\dagger = (\mathbf{B}^i)^H \mathbf{B}^i\). It can be seen that the positive sequence phasor ML estimator is absent from the MDL expression in (24). This result stems from the fact that the positive sequence appear under any hypothesis and, hence, cannot be used to distinguish between hypotheses. Therefore, the MDL criterion in (24) can be interpreted as a detector of the presence of the negative sequence phasors, with a penalty function that inhibits the number of unbalanced buses, \(|S_i|\), from becoming very large.

The MDL approach is summarized in Algorithm 1.

For the special case of voltage measurements only, \(K = M\), \(\mathbf{B} = \mathbf{I}_M\) and, thus, the state estimators from (20) and the MDL from (24) are reduced to

\[
\hat{c}_+^i = \frac{1}{1 - |\eta|^2} (z_+^i - \eta^* z_+^i) \quad \hat{c}_-^i = \frac{1}{1 - |\eta|^2} (z_-^i - \eta^* z_-^i),
\]

and

\[
T_i^{(\text{MDL})} = \frac{N}{\sigma^2(1 - |\eta|^2)} \left| \left| z_+^i - \eta^* z_+^i \right| \right|^2 - |S_i| \log N,
\]

respectively.

IV. LOW-COMPLEXITY IDENTIFICATION METHODS

Detecting multiple simultaneous imbalances in power networks using the MDL approach from Algorithm 1 is time-consuming, due to the number of hypotheses that grow exponentially with the network size. In particular, the MDL approach requires an exhaustive search and evaluation of \(T_i^{(\text{MDL})}\) over all candidate subsets of unbalanced buses. In this section, we develop two low-complexity identification methods. The first is a decentralized, local approach, in which the imbalance detection and state estimation are performed separately at each bus. The second is a computationally-efficient CS technique that is developed by using the sparsity of the negative sequence phasor vector, \(c_-^i\).

### Algorithm 1: MDL-based networked unbalanced bus identification and state estimation

**Input:** Observation vectors \(z[n], n = 0, \ldots, N - 1\), nominal frequency \(\omega_0\), and system topology \(\mathbf{B}\).

**Output:** Subset of unbalanced buses, \(S^{(i_{\text{opt}})}\), and state estimators, \(c_+^{(i_{\text{opt}})}\) and \(c_-^{(i_{\text{opt}})}\). for \(i = 0, 1, \ldots, P\) do

1. Estimate the negative-sequence vector, \(\hat{c}_-\), via (20)
2. Substitute \([S_i]\) and \(c_+\) in (24) to obtain the MDL statistic for the \(i\)th hypothesis, \(T_i^{(\text{MDL})}\)
end

Select the optimal subset, \(S^{(i_{\text{opt}})}\), by:

\[
i_{\text{opt}} = \arg \max_{i = 0, 1, \ldots, P} T_i^{(\text{MDL})}.
\]

Compute the state estimators from (20) and (21) for the selected subset of unbalanced buses, \(S^{(i_{\text{opt}})}\):

\[
\hat{c}_+^{(i_{\text{opt}})} = \frac{1}{1 - |\eta|^2} \mathbf{B}^\dagger (G_{i_{\text{opt}}} z_+ - \eta \mathbf{B}^{(i_{\text{opt}})}) \quad \hat{c}_-^{(i_{\text{opt}})} = \frac{1}{1 - |\eta|^2} \mathbf{B}^\dagger (G_{i_{\text{opt}}})\, \mathbf{B} \mathbf{B}^\dagger z_+,
\]

where \(z_+, z_-, \eta\), and \(G_i\) are defined in Subsection III-B.

---

Fig. 2. Single PMU model at the \(n\)th bus with a single voltage and \(L_m\) currents measurements.

### A. Local detection

In this subsection, the MDL approach is performed separately at each bus. Thus, the proposed local MDL approach has the advantage that it can be implemented for distributed detection, without using the central information. At the \(m\)th bus, the PMU measures synchronously the voltage at the bus and the current flows on lines incident to this bus, as presented schematically in Fig. 2. Based on the local observations, \(z_m[n], n = 0, \ldots, N - 1\), from the \(m\)th bus, we seek to identify all the observable buses. Therefore, in this case, \(K = 1\), \(\mathbf{z}_+, m \triangleq \sum_{n=0}^{N-1} e^{-j \frac{\omega_0}{\omega} \nu_{m}^T \mathbf{Y}_m[n]} z_m[n]\), \(\mathbf{z}_-, m \triangleq \frac{1}{N} \sum_{n=0}^{N-1} e^{-j \frac{\omega_0}{\omega} \nu_{m}^T \mathbf{Y}_m[n]} e_{m}^T \mathbf{Y}_m[n]\), and the local system submatrix is

\[
\mathbf{B} = \left[ \begin{array}{c} \mathbf{e}_m^T \\ \mathbf{Y}_m \end{array} \right],
\]

in which \(e_m\) is a zeros vector with one at the \(m\) row and \(\mathbf{Y}_m\) relates the current phasors to the state vector at the \(m\)th bus. By using these values in Algorithm 1, we obtain the local MDL approach.

In particular, if only voltage measurements are available, then it can be shown that the local MDL approach reduces to a local detector, which declares that the \(m\)th bus is unbalanced when

\[
T_m^{(\text{local})} \triangleq \frac{N}{\sigma^2} \left| z_{-, m} - \eta^* z_{+, m} \right|^2 > \log N (1 - |\eta|^2).
\]

The local MDL test, \(T_m^{(\text{local})}\), from (28) provides the same decision rule as the GLRT for a binary hypothesis-testing
problem with a fixed threshold. The proposed GLRT approach in (28) is the first method for local imbalances detection that is based solely on the positive sequence. In addition, by comparing (26) and (28), it can be verified that the local MDL coincides with the centralized MDL where only voltage measurements are available. Therefore, in this case, we can obtain the networked performance by applying a detector at each bus separately.

B. The POMP method

In order to use CS techniques [31], we first observe that in the measurement model from (11) the negative-sequence state vector, \( c_\gamma \), is a sparse vector. That is, \( c_\gamma \) has only a few nonzero elements that are related to the unbalanced buses. Thus, the low-complexity POMP method from [49] may be exploited to identify the support of \( c_\gamma \), as well as greedy CS algorithms with partial known support (e.g. [52]–[54]). The POMP concept is a low-complexity CS technique for estimating two vectors in a linear Gaussian model, where one of the unknown vectors is subject to a sparsity constraint. Here we adopt the POMP method to the specific model in (11).

First, we notice that, similarly to the derivation of (20) and (23), for a given sparsity level, \( s \), the ML estimator of \( c_\gamma \) is given by

\[
\hat{c}_\gamma^s = \arg \min_{c_\gamma, s.t. \|c_\gamma\|_0 = s} \| \mathbf{z} - \eta^* \mathbf{B}^\dagger \mathbf{z}_+ - \mathbf{B} c_\gamma \|^2_2.
\]

Therefore, based on our work from [49], the POMP algorithm can be applied on our problem. POMP for our model consists of two stages:

1) Imbalance identification: The OMP method iteratively finds the locations of the unbalanced buses, i.e. the support set \( S \) of the vector \( c_\gamma \), by using the projected measurements \( \tilde{\mathbf{z}} \triangleq \frac{1}{\|\mathbf{z}\|_2} (\mathbf{z} - \eta^* \mathbf{B}^\dagger \mathbf{z}_+). \) At each iteration, the OMP algorithm proceeds by finding the column of the matrix \( \mathbf{B} \) that correlates most closely to the current residual, \( \mathbf{r}^{(k)} \), where \( k \) is the iteration index. The residual \( \mathbf{r}^{(k)} \) is obtained by subtracting the contribution of the current estimate of \( c_\gamma \) from the projected measurements.

2) State estimation: The final estimation of the support set, \( \tilde{S} \), is used to calculate \( \mathbf{B}^{(i)} = \mathbf{B}_{\tilde{S}} \). Then, the state vectors, \( [c_\gamma^s]_{\tilde{S}} \) and \( [c_\gamma]_{\tilde{S}} \), are estimated by substituting \( \mathbf{B}^{(i)} = \mathbf{B}_{\tilde{S}} \) in (20) and (21), respectively.

The proposed POMP method is described in Algorithm 2.

C. Conditions for solution existence

Since we assumed that the system is perfectly observable by the installed PMUs, the matrix \( \mathbf{B} \) is a full column rank matrix and \( K + L \geq M \). In addition, the state estimators in (20) and (21) are based on the assumption that \( N > 1 \) in order to have \( \eta \neq 1 \). In general, the unknown phasor vectors, \( c_+ \) and \( c_- \), have \( 2M \) unknown complex parameters. Therefore, in order to have a unique solution, we require that our observability satisfy \( N(K + L) \geq 2M \). Together with the observability assumption, it can be verified that \( N \geq 2 \) time samples is a sufficient condition for a unique solution. Since the POMP approach requires estimation of only \( M + |S| \) unknown parameters, as well as identification of the support set, where \( |S| \) is the number of unbalanced buses, then less measurements and/or less PMUs are required. A deeper discussion for the general case, in terms of CS bounds, can be found in [55].

D. Extension to three-phase data

In current power systems, the PMUs transmits only the positive sequence, due to communication bandwidth and processing cost limitations. In the future, wide area monitoring systems may be configured for three-phase measurements. The methods proposed herein can be readily extended to the three-phase measurement case, as we now described.

Similar to (5), by substituting (1)-(3) in (9), we obtain

\[
v_{-,m}[n] = e^{j\gamma \frac{v_{n-1}^+}{\omega_0}} C_{-,m} + e^{-j\gamma \frac{v_{n-1}^+}{\omega_0}} C_{-,m}^*, \quad (31)
\]
for any \( n = 0, \ldots, N - 1 \) and any bus \( m = 1, \ldots, M \). Following (11), the model of the negative sequence at the control center is given by

\[
y^*[n] = e^{j\gamma\frac{n-\Delta}{\omega_0}} B^* c_+ + e^{-j\gamma\frac{n-\Delta}{\omega_0}} B^* c_- + w_-[n],
\]

where \( w_-[n] \) is the networked negative-sequence complex circularly symmetric Gaussian noise at time \( n \), which is assumed to be uncorrelated with \( w[n], \) \( n = 0, \ldots, N - 1 \).

If both the negative and positive sequences are available, then the joint measurement model that merges (11) and (32) is given by

\[
\tilde{z}[n] = e^{j\gamma\frac{n-\Delta}{\omega_0}} B^* c_+ + e^{-j\gamma\frac{n-\Delta}{\omega_0}} B^* c_- + w_-[n],
\]

where \( \tilde{z}[n] \triangleq [z[n]^T, y^H[n]]^T \) and \( \tilde{w}[n] \triangleq [w[n]^T, w_-^T[n]]^T \). Since the model in (33) is identical to that in (11), the proposed MDL, local MDL, and POMP can be implemented for the two-phase data by replacing the measurement vectors \( z[n] \) with \( \tilde{z}[n] \) for any \( n = 0, \ldots, N - 1 \), and by replacing the topology matrix \( B \) with \( B \).

In a similar fashion, the proposed methods can be extended to identify zero-phase imbalance scenarios by using three-phase data and obtaining the associated network measurement model as in (11) and (32).

V. SIMULATIONS

In this section, we compare the performance of five methods in three different networks. The methods are: the VUF [17], [18], which is a local, two-phase based detection method that uses the voltage measurements, the GLRT from [24], which is a local, three-phase based detection method that uses the voltage measurements, and the proposed MDL, local MDL, and POMP methods. The networks are for a small two-port π-model network, the IEEE 14-bus, and IEEE 118-bus. The MDL method from Algorithm 1 is applicable only for the first example, due to its computational complexity. The centralized POMP method is implemented with a threshold proportional to \( \tau = |B|^2 \sigma^2 \). The commonly-used VUF statistic is defined as the voltage magnitude ratio of the negative to the positive sequences. In the following simulations, the VUF at the \( m \)th bus is defined as

\[
T_{VUF,m} = \frac{\sum_{n=0}^{N-1} |v_{-,m}[n]|}{\sum_{n=0}^{N-1} |v_{+,m}[n]|}.
\]

For an unbalanced bus, \( T_{VUF,m} \) should be close to 1. Therefore, in the following simulations the decision of the VUF detector is based on comparing \( T_{VUF,m} - 1 \) with a threshold. In all simulations, we assume a nominal grid frequency of \( \omega_0 = 2\pi/60 \) and an unknown frequency deviation of \( \Delta = 0.2\pi \), which is estimated via (22). The performance is evaluated using 10,000 Monte-Carlo simulations.

A. Two-port π-model

We assume the two-port π-model (see e.g. [8] and p. 151 in [1]). We consider two PMUs, one at each end of the transmission line, with a sampling rate of \( N = 48 \) samples per cycle of \( \omega_0 \). The measurement equation is given by (10), where \( v_+[n] = [v_{1,+}[n], v_{2,+}[n]]^T \) and

\[
B = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
y_{12} + y_{10} & -y_{12} \\
-y_{12} & y_{20}
\end{bmatrix},
\]

in which \( y_{12}, y_{10}, y_{20}, y_{12} \) represent the corresponding positive sequence series admittances. In this case, there are \( P = 4 \) hypotheses with balanced/unbalanced conditions for each bus.

A single-phase voltage magnitude imbalance is implemented in bus 2 by setting the parameters as defined in Table I, where the voltage magnitudes are considered to be per unit (p.u.). The admittance parameters are set to \( y_{12} = 0.1 \), \( y_{10} = 0.5 \), and \( y_{20} = y_{12} \). In Fig 3, the probability of error of the different approaches is presented versus \( \beta \) for a noise variance of \( \sigma^2 = -5,0dB \) and \( N = 48 \). The probability of error is the total probability that the methods incorrectly detected that bus 2 is balanced and/or that bus 2.

<table>
<thead>
<tr>
<th>Voltage magnitudes</th>
<th>Voltage phases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus 1 ( V_a = V_b = V_c = 1 ) ( (\phi_1, \phi_2, \phi_3) = (-\pi, \pi, \pi) )</td>
<td>( (\phi_1, \phi_2, \phi_3) = (-\pi, \pi, \pi) )</td>
</tr>
<tr>
<td>Bus 2 ( V_a = V_c = 1.25 ), ( V_b = \beta V_a ), ( \beta &gt; 1 ) ( (\phi_1, \phi_2, \phi_3) = (-\pi, \pi, \pi) )</td>
<td>( (\phi_1, \phi_2, \phi_3) = (-\pi, \pi, \pi) )</td>
</tr>
</tbody>
</table>

**TABLE I**

Example 1 (Two-port π-model): Simulation parameters

![Fig. 3. Two-port π-model: Probability of error for two-port π-model with \( \sigma^2 = -5,0dB \), \( N = 48 \), and single-phase voltage magnitude imbalance at bus 2.](image-url)
taken from [56]. The PMUs are located at buses 2, 6, 7, 9 and measure bus voltages and currents in lines connected to these buses, which results in an observable system [44]. This setting results in a total of 18 sensor measurements and 14 unknown positive sequence phasors. At each Monte Carlo run, PMU measurements were generated as an independent identically distributed (i.i.d.) Gaussian random vector with the mean equal to the operating point given in the IEEE 14-bus data [56] and covariance matrix $\sigma^2 I_{133}$. Similar to [2], loading imbalances are found at the load buses 4, 5 and 9. To generate single-phase imbalances, we multiply the amplitude of phase $c$ load by $\beta > 1$ at the unbalanced buses. The local MDL method is performed based on the single PMU located at bus 7. Thus, the local method is only able to identify the imbalance condition of buses 2, 6, 7, 9.

Figure 4 shows the Receiver Operating Characteristic (ROC) curves of the VUF, GLRT, local MDL, and POMP methods for the detection of imbalance at bus 9 for two scenarios: 1) $\beta = 3.75$ and $N = 24$; and 2) $\beta = 1.1$ and $N = 12$. It can be seen that the POMP and local MDL approaches significantly outperform the VUF and GLRT, where for small values of $\beta$ the imbalance is undetectable by the VUF and GLRT. In Fig. 5, the probability of detection of all the unbalanced buses, 4, 5, and 9, is presented versus $\beta$ for $N = 10$ and SNR $= 10db$, where SNR $\triangleq \frac{N}{\sigma^2}$. The probability of false alarm, is around 5% for all methods. It can be seen that the detection probability is higher for the POMP than for the local MDL method at buses 4 and 9, and than for the VUF and GLRT methods at bus 9. The reason is that the local MDL approach is based only on the single-bus PMU measurements, which results in lower accuracy compared with centralized approaches. The VUF and GLRT use only voltage measurements, which results in a lower detection probability compared with the single-phase approaches, POMP and local MDL. Therefore, it can be concluded that the use of both current and voltage single-phase measurements is better than the use of two/three-phase voltage-only data. Moreover, imbalance at bus 5 is undetectable by the local methods, since this bus is unobservable by the PMU located at buses 7 or 9. That is, the POMP approach is the only method that achieves full observability of the system.

C. Case study: Loading imbalances in large-scale networks

The purpose of this experiment is to investigate the performance of the proposed algorithm when applied to large systems. In particular, we study loading imbalances in the IEEE 118-bus system, where the system parameters are taken from [56] and the PMUs are located according to the “base case state” outlined in Table V in [57], which results in an observable system [44], [57]. This setting results in a total of 133 sensor measurements and 118 unknown positive sequence phasors. At each Monte Carlo run, PMU measurements were generated as an i.i.d. Gaussian random vector with the mean equal to the operating point given in the IEEE 118-bus data [56] and the covariance matrix $\sigma^2 I_{133}$. Loading single-phase imbalances are implemented at load buses, that are randomly chosen to be 3, 9, 33, 39, 51, 57, 75, 93, 96, 98, with $\beta > 1$ at the unbalanced buses. The local MDL method is performed, based on the single PMU located at bus 9. Thus, the local method is only able to identify the imbalance condition of buses 8, 9 and 10.

Figure 6 shows the ROC curves of the VUF, GLRT, local MDL, and POMP methods, where bus 9 is unbalanced for two scenarios: 1) $\beta = 3.75$ and $N = 24$ (solid line); and 2) $\beta = 1.1$ and $N = 12$ (dashed line). It can be seen that the POMP and local MDL approaches significantly outperform the VUF and the GLRT methods, for which at small values of $\beta$ the imbalance is undetectable. The behavior of MDL can be explained by the fact that this criterion has been shown to be inconsistent when the variance of the noise tends to zero [58]. In Fig. 5, the probability of detection of the unbalanced buses is presented versus $\beta$ for $N = 10$. The probability of false alarm, is around 10% for all methods. Evidently, the detection probability is higher for POMP than for the local MDL method at bus 9, and than for the VUF and GLRT at bus 9. This effect is expected to be higher for a larger network. The reason is that the local MDL approach is based only on the local single-bus PMU measurements, which results in lower accuracy compared with centralized approaches.

Comparison between the results of the IEEE 14-bus system in Figs. 4 and 5 and those of the IEEE 118-bus system in Figs. 6 and 7, indicates that the VUF and local MDL techniques have better performance for the smaller 14-bus system, the...
GLRT method has better performance for the large 118-bus system, and the POMP algorithm performs well for both the 14-bus and 118-bus systems.

Fig. 6. Loading imbalances for IEEE 118-bus system: The ROC curves of VUF, GLRT, local MDL, and POMP methods versus the number of unbalanced load buses, which is equivalent to the sparsity level, is presented for any problem dimensions, the average processing period, "runtime", was evaluated by running the algorithm using Matlab on an Intel Core(TM) i7-5820K CPU computer, 3.3 GHz. Figure 9 shows the runtime of the POMP method as a function of the percent of unbalanced load buses in the system for 14-bus and 118-bus systems and $N = 12, 24$ samples. It can be seen that the runtime is higher for the 118-bus system than for the 14-bus system, as expected, but is has the same order. In addition, the runtime increases approximately linearly with the problem dimensions. It should be noted that POMP is the only centralized method that is not restricted to very small problem dimensions.

In this paper, we proposed two methods to identify and localize imbalances occurring at any location in the power grid and, simultaneously, estimate the states, based on single-phase PMU measurements. The single-phase data affords low communication and processing costs, compared to three-phase data processing. The proposed MDL and POMP approaches use central information from the entire network, where the POMP method has reduced computational complexity. The difference between the centralized and the distributed implementations is that the former performs the state estimation and imbalance detection steps over the complete data set and, thus, is more accurate, whereas in the latter each bus runs these steps on its own local data. Simulation results confirm that the proposed approaches are feasible and that the centralized detection methods, MDL and POMP, display superior performance.

VI. CONCLUSION

In this paper, we proposed two methods to identify and localize imbalances occurring at any location in the power grid and, simultaneously, estimate the states, based on single-phase PMU measurements. The single-phase data affords low communication and processing costs, compared to three-phase data processing. The proposed MDL and POMP approaches use central information from the entire network, where the POMP method has reduced computational complexity. The difference between the centralized and the distributed implementations is that the former performs the state estimation and imbalance detection steps over the complete data set and, thus, is more accurate, whereas in the latter each bus runs these steps on its own local data. Simulation results confirm that the proposed approaches are feasible and that the centralized detection methods, MDL and POMP, display superior performance.
compared to the local MDL method. For small-scale networks, it is shown that the performance of POMP convergences to that of the asymptotically-consistent MDL approach. However, the centralized MDL, which has the lowest probability of error, cannot be implemented on large-scale networks. Thus, the proposed POMP method provides the best performance, in terms of high probability of detection and low runtime for large-scale networks. The local MDL method achieves better performance compared with local methods, i.e. GLRT and VUF, and the performance of the local MDL converges to that of the POMP algorithm for high SNR and/or a large number of measurements.

REFERENCES

[42] A. G. Phadke, J. S. Thorp, and M. G. Adamiak, “A new measurement technique for tracking voltage phasors, local system frequency, and rate


