A Sparse Representation Approach to Online Estimation of Power System Distribution Factors

Yu Christine Chen, *Student Member, IEEE*, Alejandro D. Domínguez-García, *Member, IEEE*, and Peter W. Sauer, *Life Fellow, IEEE*

***Abstract*—In this paper, we propose a method to compute linear sensitivity distribution factors (DFs) in near real-time without relying on a power flow model of the system. Specifically, we compute the injection shift factors (ISFs) of a particular line of interest with respect to active power injections at all buses (all other DFs can be determined from ISFs). The proposed ISF estimation method relies on the solution of an underdetermined system of linear equations that arise from high-frequency synchronized measurements obtained from phasor measurement units. We exploit a sparse representation (i.e., one in which many elements are zero) of the vector of desired ISFs via rearrangement by electrical distance and an appropriately chosen linear transformation, and cast the estimation problem into a sparse vector recovery problem. As we illustrate through case studies, the proposed approach provides accurate DF estimates with fewer sets of synchronized measurements than earlier approaches that rely on the solution of an overdetermined system of equations via the least-squares errors estimation method.**

I. Introduction

Existing tools for online power system operational reliability and health monitoring rely on a system model obtained offline, constructed from the transmission network, line parameters, and historical and forecasted power generation and demand [1]. For example, real-time contingency analysis (RTCA) determines whether or not the system will meet operational reliability requirements in the case of outage in any one particular system asset (e.g., a transmission line or generator), a condition known as N-1 security [2]. Using an up-to-date system model, operators can perform RTCA by repeated solutions of nonlinear power flow problems. An alternative to repeated computations of the nonlinear power flow solution is to use linear sensitivity distribution factors (DFs), such as injection shift factors (ISFs) and line outage distribution factors (LODFs), to predict the effect of a hypothetical operating point change on the system [3]. In the case that N-1 security criterion is not satisfied, DFs can also be used to re-dispatch generators to relieve transmission line loading [4]. Conventionally, ISFs are obtained offline, for some nominal operating point, from a power flow model—all other DFs can be obtained from ISFs; hence we refer to analysis tools that rely on DFs computed offline using the power flow model as "model-based". In this paper, we propose a method to estimate ISFs in near real-time using PMU measurements without relying on a power flow model of the system; hence we refer to analysis tools that rely on DFs obtained this way as "measurement-based".

Model-based online analysis tools are not ideal since their results rely on an accurate model that reflects up-to-date network topology and parameters, which may not be available due to erroneous records or telemetry. The results of model-based analyses lack the flexibility of adapting to unexpected changes in network topology or variations in generation and load, all of which can affect the power flow solution significantly, and in turn any corrective actions established based on them. For example, in the 2011 San Diego blackout, operators could not detect that certain lines were overloaded or close to being overloaded because the network configuration in the model was not up-to-date [1]. Thus, traditional model-based techniques may no longer satisfy the needs of monitoring and protection tasks; therefore there exists an impetus to shift away from model-based analyses and to develop measurement-based power system online analysis tools that are adaptive to changes in operating point and topology [1], [2]. With respect to this, phasor measurement units (PMUs) are an enabling technology for the development of such measurement-based online analysis tools. For example, the authors in [5]–[7] proposed the use of PMU measurements (along with a model of the power system) for detecting line outages.

In accordance with the vision described above to shift away from model-based analyses, in [8], [9], we proposed a measurement-based method to estimate DFs by finding the solution of a system of linear equations formulated using active power bus injection and line flow data obtained from PMU measurements. Specifically, in [9], we considered an overdetermined system, with more equations than unknown ISFs, and obtained the solution via linear least-squares errors (LSE) estimation. While the method is shown to accurately compute ISFs, even in the presence of undetected system topology and operating point changes, the LSE estimation problem formulation necessitates at least as many sets of synchronized measurements as unknown ISFs. For a large power system, such a restriction may be ill-advised in, e.g., RTCA, since power systems are constantly undergoing changes and operators often need to quickly determine whether or not the current system is secure. To address this issue, we propose an accurate and efficient method to recover the ISF solution using fewer sets of measurements than unknown ISFs. To this end, we exploit a sparse representation (i.e., one in which many elements are zero) of the vector of desired ISFs, solve for the transformed sparse representation, and finally compute

The authors are with the Department of Electrical and Computer Engineering of the University of Illinois at Urbana-Champaign, Urbana, IL, 61801, USA. E-mail: {chen267, aledan, psauer}@illinois.edu.

The work described in this paper was made possible by funding provided by the U.S. Department of Energy under the Consortium for Electric Reliability Technology Solutions (CERTS); and by the Natural Sciences and Engineering Research Council of Canada under its Postgraduate Scholarship Program.

the original ISFs by applying the inverse transformation. The work presented in this paper, for the most part, assumes that all buses within the monitored region are equipped with PMUs. Admittedly, present-day power systems are still far from having such a rich set of available phasor measurement devices; however, incentives to invest in the deployment of such measurement infrastructure are driven by preliminary demonstrations of its potential benefits in monitoring, protection, and control capabilities (see, e.g., [10]). Moreover, today, in addition to PMU installations, synchronous phasor measurement capabilities are available as standard features in many protective relays, meters, and recorders [11].

This paper builds on preliminary work reported in [12], providing extensions in several directions. First, we introduce an alternative method to sparsify ISFs based on electrical distance using only the network Z-bus matrix. This method is advantageous over the one proposed in [12] in that, like the measurement-based ISF computation, the electrical distance is independent of slack bus location. Second, we explore two algorithms to recover the sparse ISF representation, and compare their performances in terms of computation time and accuracy. Further, we deploy the method to two large case studies, showcasing its scalability. Finally, the requirement that active power injection data must be available at all buses is relaxed so that only a subset of them are needed.

The remainder of this paper is organized as follows. Section II describes the measurement-based ISF estimation problem setup, and the LSE-based solution proposed in [8]. Section III introduces the sparse representation solution approach that we propose in this paper, illustrating it via an example involving the IEEE 14-bus system. In Section IV, we formulate two algorithms to solve the ISF estimation problem using fewer sets of synchronized measurements than number of unknowns. In Section V, we present case studies involving the IEEE 300-bus and the Polish 2383-bus systems, and highlight the effectiveness of the proposed ISF estimation method. We provide a variation that uses a subset of measurements to estimate ISFs in Section VI. Finally, in Section VII, we offer concluding remarks and directions for future research.

II. Preliminaries

Distribution factors are linearized sensitivities used in, e.g., contingency analysis and transmission loading relief [4]. A key distribution factor is the injection shift factor (ISF), which quantifies the redistribution of power through each transmission line in a power system following a change in generation or load on a particular bus in the system. In essence, the ISF captures the sensitivity of the flow through a line with respect to changes in generation or load. In this section, we provide the mathematical definition of the ISF and introduce relevant notation. We also summarize the ISF estimation approach proposed in [8], and conclude by highlighting its main drawback.

A. Injection Shift Factor Definition

The ISF of line $L_{k\text{-}l}$ (assume positive active power flow from bus k to l) with respect to bus i, denoted by $\Psi^i_{k\text{-}l}$, is a linear approximation of the sensitivity of the active power flow in line $L_{k\text{-}l}$ with respect to the active power injection at bus i, with the slack bus defined and all other quantities constant. Then, based on the definition,

$$\Psi^i_{k\text{-}l} := \frac{\partial P_{k\text{-}l}}{\partial P_i}. \tag{1}$$

Let $P_i(t)$ and $P_i(t+\Delta t)$, respectively, denote the active power injection at bus i at times t and $t+\Delta t$, $\Delta t > 0$ and small. Define $\Delta P_i(t) = P_i(t+\Delta t) - P_i(t)$, and denote by $\Delta P^i_{k\text{-}l}(t)$ the change in line $L_{k\text{-}l}$ active power flow resulting from $\Delta P_i(t)$. Then, it follows from (1) that

$$\Psi^i_{k\text{-}l} \approx \frac{\Delta P^i_{k\text{-}l}(t)}{\Delta P_i(t)}. \tag{2}$$

B. Injection Shift Factor Estimation

In order to obtain $\Psi^i_{k\text{-}l}$, we need $\Delta P^i_{k\text{-}l}(t)$, which is not readily available from PMU measurements. We assume that the net variation in active power through line $L_{k\text{-}l}$, denoted by $\Delta P_{k\text{-}l}(t)$, however, is available from PMU measurements. We express this net variation as the sum of active power variations in line $L_{k\text{-}l}$ due to active power injection variations at each bus i:

$$\Delta P_{k\text{-}l}(t) = \Delta P^1_{k\text{-}l}(t) + \cdots + \Delta P^n_{k\text{-}l}(t). \tag{3}$$

Equivalently, by substituting (2) into (3), we can rewrite (3) as

$$\Delta P_{k\text{-}l}(t) \approx \Delta P_1(t)\Psi^1_{k\text{-}l} + \cdots + \Delta P_n(t)\Psi^n_{k\text{-}l},$$

where $\Psi^i_{k\text{-}l} \approx \frac{\Delta P^i_{k\text{-}l}}{\Delta P_i}$, $i = 1,\ldots,n$. Suppose $m+1$ sets of synchronized measurements are available. Let $\Delta P_i[j] = P_i((j+1)\Delta t) - P_i(j\Delta t)$, and $\Delta P_{k\text{-}l}[j] = P_{k\text{-}l}((j+1)\Delta t) - P_{k\text{-}l}(j\Delta t)$, $j = 1,\ldots,m$; and define $\Delta P_{k\text{-}l} = [\Delta P_{k\text{-}l}[1],\ldots,\Delta P_{k\text{-}l}[j],\ldots,\Delta P_{k\text{-}l}[m]]^T$, and $\Delta P_i = [\Delta P_i[1],\ldots,\Delta P_i[j],\ldots,\Delta P_i[m]]^T$. Let $\Psi_{k\text{-}l} = [\Psi^1_{k\text{-}l},\ldots,\Psi^i_{k\text{-}l},\ldots,\Psi^n_{k\text{-}l}]^T$; then, it follows that

$$\Delta P_{k\text{-}l} \approx \begin{bmatrix}\Delta P_1 & \cdots & \Delta P_i & \cdots & \Delta P_n\end{bmatrix}\Psi_{k\text{-}l}. \tag{4}$$

For ease of notation, let ΔP represent the $m \times n$ matrix $[\Delta P_1,\ldots,\Delta P_i,\ldots,\Delta P_n]$; then, the system in (4) can be compactly written as

$$\Delta P_{k\text{-}l} \approx \Delta P \Psi_{k\text{-}l}. \tag{5}$$

As we proposed in [8], if $m \geq n$, then (5) is an overdetermined system, and we can solve for $\Psi_{k\text{-}l}$ via LSE estimation as follows:

$$\hat{\Psi}_{k\text{-}l} = (\Delta P^T \Delta P)^{-1}\Delta P^T \Delta P_{k\text{-}l}. \tag{6}$$

However, for a large system with many buses, it may not be prudent to require such a large number of datasets before an estimate can be computed. Further, the adaptability of the measurement-based approach to system changes would be improved if fewer sets of data are required. Thus, in Sections III–V, we focus on solving for $\Psi_{k\text{-}l}$ in (5) when $m < n$, i.e., obtaining a solution when (5) is an underdetermined system of equations. In Section VI, we discuss a modification to the proposed method and solve for $\Psi_{k\text{-}l}$ with active power injection data available at fewer than n buses.

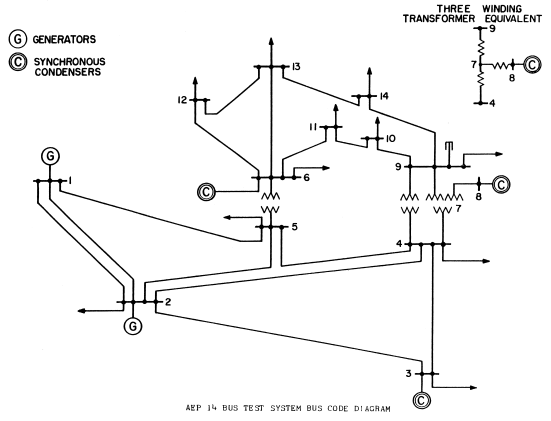


Fig. 1: Network topology for IEEE 14-bus system [16].

III. A Sparse Representation Approach

The computation approach presented in this paper is inspired by the field of compressive sensing (CS) (see, e.g., [13]), and its applications to image processing, where a typical problem is to compress a large image (i.e., to reduce irrelevant or redundant image data in order to store or transmit the image efficiently), and subsequently reconstruct the image from its compressed representation. [Recently, CS ideas have been applied to the identification of multiple line outages in power systems [6].] CS theory asserts that, by exploiting their sparsity, certain classes of signals can be recovered from fewer samples or measurements than those needed in traditional methods such as LSE estimation (see, e.g., [14], [15]). Specifically, the problem of recovering a sparse signal can be cast as one where the objective is to minimize the l_0-norm[1] of the signal to be recovered.

A. Sparsifying the Vector of Injection Shift Factors

In our setting, the signal of interest, the ISF vector $\Psi_{k\text{-}l}$, is not necessarily sparse; therefore, we search for an appropriate linear coordinate transformation M, such that $c_{k\text{-}l} = M\Psi_{k\text{-}l}$, where $c_{k\text{-}l}$ is a sparse vector, i.e., one in which many elements are zero. While there may exist numerous possible transformations that achieve the aforementioned objective, in this work, we use an intuitive approach based on electrical distances.

In order to compute $\Psi_{k\text{-}l}$ with $m < n$ measurements, we first reorder its entries by the electrical distance of each bus to the line $L_{k\text{-}l}$, a main advantage of which is its independence from the slack bus location. Then, we apply a *difference transformation* (to be defined below) to the reordered signal. This leads to a sparse representation of the ISF vector, which allows us to cast the ISF estimation problem into a sparse vector recovery problem. Once this problem is solved, we apply the inverse difference transformation to the resulting sparse vector estimate to obtain an estimate of the ISF vector.

1) Reordering by electrical distance: We assume that, prior to online estimation, we have a base-case system model that consists of all relevant nominal system topology, parameter, and operating point information. Using this model, we perform a one-time computation of the electrical distance from each bus to line $L_{k\text{-}l}$. The derivation below is provided in [17], where the concept of electrical distances was used to allocate the cost of transmission to generators and loads.

Denote the bus admittance matrix by[2] $\mathbf{Y} = [\mathbf{Y}_{ik}]$. Then, apply Kirchhoff's current law at each bus and express the current injected into each bus i from the ground node as

$$\mathbf{I}_i = \sum_{k=1}^{n} \mathbf{Y}_{ik}\mathbf{V}_k, \tag{7}$$

where $\mathbf{V}_k$, $k = 1, \ldots, n$, is the voltage at node k. Let $\mathbf{Z} = [\mathbf{Z}_{ki}] = \mathbf{Y}^{-1}$, then the voltage at node k can be expressed as

$$\mathbf{V}_k = \sum_{i=1}^{n} \mathbf{Z}_{ki}\mathbf{I}_i. \tag{8}$$

On the other hand, the current through the line connected from node k to l is

$$\mathbf{I}_{k\text{-}l} = (\mathbf{V}_k - \mathbf{V}_l)\mathbf{y}_{k\text{-}l} + \mathbf{V}_k\mathbf{y}^{\text{sh}}_{k\text{-}l}, \tag{9}$$

where $\mathbf{y}_{k\text{-}l}$ is the admittance of the line connecting nodes k and l, and $\mathbf{y}^{\text{sh}}_{k\text{-}l}$ is the shunt admittance at node k. Substituting (8) into (9), we obtain

$$\mathbf{I}_{k\text{-}l} = \left(\sum_{i=1}^{n} \mathbf{Z}_{ki}\mathbf{I}_i - \sum_{i=1}^{n} \mathbf{Z}_{li}\mathbf{I}_i\right)\mathbf{y}_{k\text{-}l} + \sum_{i=1}^{n} \mathbf{Z}_{ki}\mathbf{I}_i\mathbf{y}^{\text{sh}}_{k\text{-}l},$$

and after rearranging, we obtain

$$\mathbf{I}_{k\text{-}l} = \sum_{i=1}^{n} \left[(\mathbf{Z}_{ki} - \mathbf{Z}_{li})\,\mathbf{y}_{k\text{-}l} + \mathbf{Z}_{ki}\mathbf{y}^{\text{sh}}_{k\text{-}l}\right]\mathbf{I}_i.$$

Then, the electrical distance between bus i and line $L_{k\text{-}l}$ is given by [17]:

$$\mathbf{a}^i_{k\text{-}l} = (\mathbf{Z}_{ki} - \mathbf{Z}_{li})\,\mathbf{y}_{k\text{-}l} + \mathbf{Z}_{ki}\mathbf{y}^{sh}_{k\text{-}l}. \tag{10}$$

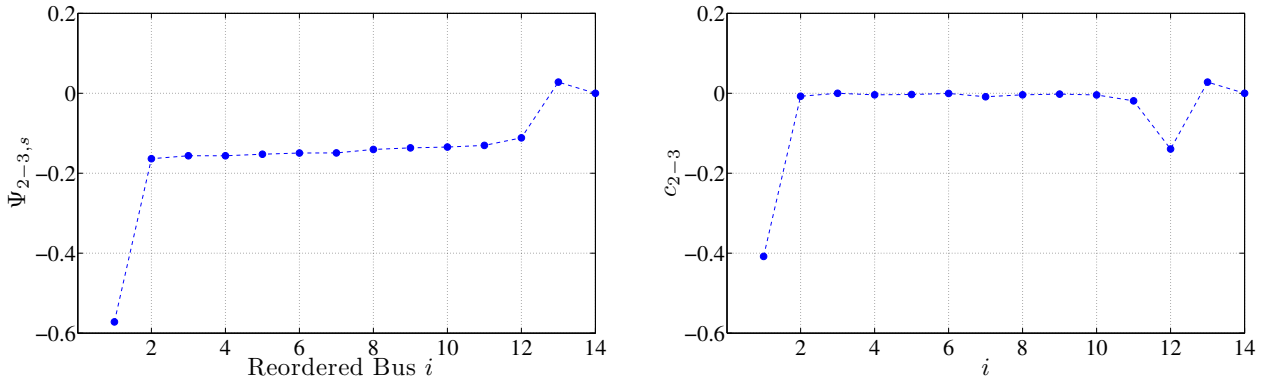
We sort the vector $\Psi_{k\text{-}l}$ according to the electrical distance as defined in (10), and denote the rearranged vector as $\Psi_{k\text{-}l,s}$, i.e.,

$$\Psi_{k\text{-}l,s} = [\Psi^1_{k\text{-}l,s}, \ldots, \Psi^i_{k\text{-}l,s}, \ldots, \Psi^n_{k\text{-}l,s}]^T,$$

such that index $i \leq j$ if $|\mathbf{a}^i_{k\text{-}l}| \geq |\mathbf{a}^j_{k\text{-}l}|$. The intuition behind this particular approach is as follows: the injections at buses that are electrically far away from line $L_{k\text{-}l}$ often have little effect on the active power flow through line $L_{k\text{-}l}$ as compared to nearer ones. Hence, after reordering of the nodes based on electrical distance, we assume that the sorted signal, $\Psi_{k\text{-}l,s}$, is characterized by smooth segments separated by sporadic jumps. Therefore, the difference between consecutive elements is likely small; as detailed below, this is the premise upon which we define a linear transformation that results in a sparse representation of $\Psi_{k\text{-}l,s}$.

[1] For a vector with finite support, the l_0-norm is defined as the number of its entries that are nonzero (see, e.g., [13]).

[2] Bolded symbols denote complex-valued quantities.



(a) Reordered ISFs but not sparse.

(b) Sparse signal after difference transformation.

Fig. 2: IEEE 14-bus system: ISFs of line $L_{2\text{-}3}$ with respect to each node pre- and post-transformation.

2) Applying a difference transformation: This transformation is defined via the difference between consecutive elements of $\Psi_{k\text{-}l,s}$:

$$c^i_{k\text{-}l} = \Psi^i_{k\text{-}l,s} - \Psi^{i+1}_{k\text{-}l,s}, \quad c^n_{k\text{-}l} = \Psi^n_{k\text{-}l,s}. \tag{11}$$

In matrix form, the difference transformation, defined in (11), can be written as

$$c_{k\text{-}l} = M\Psi_{k\text{-}l,s},$$

where $M = [m_{ij}]$, with $m_{ij} = 1$ if $j = i$, $m_{ij} = -1$ if $j = i+1$, and $m_{ij} = 0$, otherwise. Let $\hat{\Psi}_{k\text{-}l,s}$ denote the estimate for $\Psi_{k\text{-}l,s}$. If the difference transformation in (11), indeed, results in $c_{k\text{-}l}$ being sparse, we can then exploit ideas from CS to compute the estimate $\hat{c}_{k\text{-}l}$ using $m < n$ measurement data sets, and then apply the inverse difference transformation, M^{-1}, to recover $\hat{\Psi}_{k\text{-}l,s}$ as

$$\hat{\Psi}_{k\text{-}l,s} = M^{-1}\hat{c}_{k\text{-}l}. \tag{12}$$

Next, we illustrate the effectiveness of the difference transformation in (11), in conjunction with the electrical distance-based reordering, for sparsifying an ISF vector.

Example 1 (IEEE 14-Bus System): Consider the IEEE 14-bus system (see, e.g., [16]), the one-line diagram for which is shown in Fig. 1. We compute the model-based linear sensitivity ISF vector of line $L_{2\text{-}3}$, $\Psi_{2\text{-}3}$, using the partial derivative definition in (2) (see Table I, column 2) by linearizing the nonlinear AC power flow equations. This is the benchmark value to which we compare any estimation results in the remainder of the paper.

After reordering by electrical distance, we plot these model-based ISFs in Fig. 2a and list them in column 3 of Table I. While the rearranged signal is fairly smooth except at buses 1, 12, and 13, this vector has only one zero element (at the reordered bus 14). In order to get a sparse representation, we apply the difference transformation defined in (11) to the signal shown in Fig. 2a, and obtain the signal depicted in Fig. 2b (also recorded in Table I, column 4), which contains many zero or near-zero elements with the same sharp edges as in Fig. 2a. We note that the resulting post-transformation vector $c_{k\text{-}l}$ is approximately sparse, i.e., there may be many negligible near-zero (instead of exactly-zero) elements, as shown in Fig. 2b. In this regard, CS has been widely and successfully applied to the compression and recovery of large images, which may only admit approximately sparse representations (see, e.g., [18]). In Section IV, we will describe algorithms that can closely estimate approximately sparse signals (see, e.g., [15]). Additionally, through examples and case studies in Sections IV–V, we will show that approximate sparsity is sufficient to recover ISFs. ■

3) On the effect of model error: The electrical distance-based reordering described in Section III-A1 is the only component in the proposed method that relies on an offline model of the system. Thus, we comment on the effect of model error on the sparsity of the resulting transformed ISF vector $c_{k\text{-}l}$. First, we note that the electrical distance metric depends only on the system topology and line and shunt admittance values, not the load or generation levels. Therefore, we focus on the effect of topology errors, which can arise either due to base-case model error or contingencies. Small deviations away from the base-case topology, such as the outage of one or (at most) two lines, do not drastically affect the resulting transformed ISF vector sparsity, especially for large-scale systems. Next, we illustrate this with an example involving the IEEE 14-bus test system.

Example 2 (IEEE 14-Bus System): We consider the same system as in Example 1. We compute the model-based linear sensitivities, $\Psi_{2\text{-}3}$, for each case with undetected outage of any one line in the system (with the exception of line $L_{2\text{-}3}$ itself and line $L_{7\text{-}8}$, which islands the system). For each of these scenarios, we observe the corresponding $\Psi_{2\text{-}3,s}$ using the index order obtained in Example 1 and note that the sparsity pattern is preserved except in the case of undetected $L_{1\text{-}2}$ outage. ■

4) On managing bad data: PMU data may be corrupted by random errors that arise from equipment limitations in the measurement and communication devices [19]. Some grossly inadequate data, such as (i) negative voltage magnitudes, (ii) values that are orders of magnitude too large or too small, and (iii) vastly different currents in and out of a bus, can be removed prior to ISF computation based on plausibility checks [19]. In this paper, we assume that standard plausibility tests were applied to the PMU measurements before being passed to the proposed algorithms for ISF estimation. Since the

TABLE I: Comparison of $\hat{c}_{2\text{-}3}$ for the IEEE 14-bus system obtained through Examples 1-3.

i	$\Psi^i_{2\text{-}3}$	$\Psi^i_{2\text{-}3,s}$	$c^i_{2\text{-}3}$	$\hat{c}^i_{2\text{-}3}$ (via Algorithm 1)			$\hat{c}^i_{2\text{-}3}$ (via Algorithm 2)			$\hat{c}^i_{2\text{-}3}$ (via LSE)	
				$m=7$	$m=10$	$m=13$	$m=7$	$m=10$	$m=13$	$m=14$	$m=20$
1	0	-0.5719	-0.4082	-0.2916	-0.4178	-0.4094	-0.3924	-0.4154	-0.4111	-0.4124	-0.4080
2	0.0279	-0.1637	-0.0074	-0.2286	0	-0.0051	-0.0003	-0.0010	-0.0037	-0.0020	-0.0069
3	-0.5719	-0.1563	0	0.1971	-0.0085	-0.0005	0.0001	-0.0002	-0.0012	-0.0025	0.0001
4	-0.1637	-0.1563	-0.0039	-0.0202	0	-0.0023	-0.0001	-0.0051	-0.0023	-0.0009	-0.0041
5	-0.1115	-0.1524	-0.0030	-0.0040	0.0010	-0.0063	-0.0000	-0.0036	-0.0054	-0.0070	-0.0013
6	-0.1304	-0.1494	-0.0004	0	-0.0049	0.0010	-0.0108	-0.0012	0.0002	0.0002	-0.0011
7	-0.1563	-0.1491	-0.0087	0	-0.0011	-0.0046	-0.0015	-0.0013	-0.0036	0.0002	-0.0118
8	-0.1563	-0.1404	-0.0039	-0.0082	0	-0.0089	-0.0123	-0.0077	-0.0084	-0.0104	-0.0021
9	-0.1524	-0.1366	-0.0021	0	-0.0130	0	0.0000	-0.0001	-0.0011	-0.0018	-0.0011
10	-0.1494	-0.1345	-0.0041	0	0	-0.0030	-0.0003	-0.0032	-0.0027	-0.0022	-0.0050
11	-0.1404	-0.1304	-0.0189	-0.1587	-0.0219	-0.0175	-0.0244	-0.0163	-0.0174	-0.0169	-0.0189
12	-0.1345	-0.1115	-0.1394	0	-0.1414	-0.1408	-0.1328	-0.1416	-0.1409	-0.1417	-0.1393
13	-0.1366	0.0279	0.0279	0	0.0315	0.0270	0.0290	0.0252	0.0269	0.0273	0.0279
14	-0.1491	0	0	0	0.0002	-0.0010	-0.0072	-0.0016	-0.0011	-0.0017	0.0003
			$\|\|\hat{c}_{2\text{-}3} - c_{2\text{-}3}\|\|_2$	0.3761	0.0225	0.0087	0.0265	0.0138	0.0094	0.0147	0.0043

proposed method exploits the difference between consecutive measurements, it is immune against errors that involve constant offsets. Moreover, the effect of bad data can be reduced or eliminated by conducting estimation over a shortened sliding window in time so that any erroneous data eventually become ineffectual as more recent measurements are acquired.

B. Sparse ISF Vector Recovery Problem

Through reordering and transforming the original ISF vector $\Psi_{k\text{-}l}$, we assume the post-transformation signal $c_{k\text{-}l}$ to be sparse (this was illustrated in Example 1). Since the elements of $\Psi_{k\text{-}l}$ have been sorted by electrical distance, we also rearrange the columns of ΔP accordingly, and denote this reordered matrix as ΔP_s. With this, we transform the original problem of estimating $\Psi_{k\text{-}l}$ such that (5) is satisfied, to the problem of estimating $c_{k\text{-}l}$ such that

$$\Delta P_{k\text{-}l} = \Phi c_{k\text{-}l} \tag{13}$$

is satisfied, where $\Phi = \Delta P_s M^{-1}$ is the so-called measurement matrix. Since $c_{k\text{-}l}$ is assumed to be a sparse vector, we can cast the ISF vector estimation problem as an optimization program

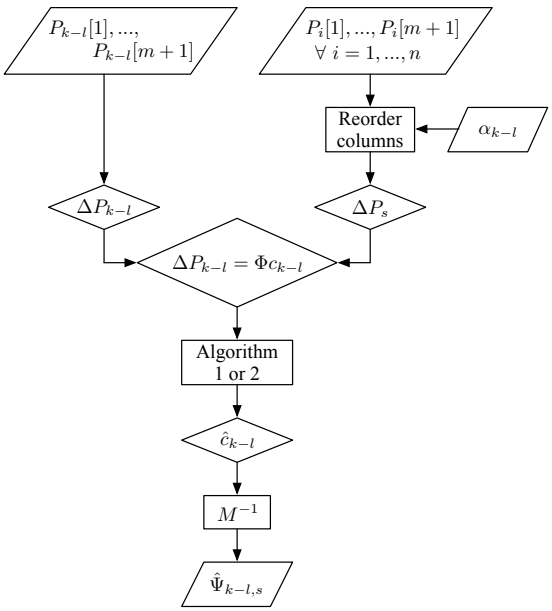


Fig. 3: Proposed ISF estimation procedure.

where the objective is to minimize $||c_{k\text{-}l}||_0$—the number of nonzero elements in $c_{k\text{-}l}$:

$$\begin{aligned} &\min_{c_{k\text{-}l}} ||c_{k\text{-}l}||_0 \\ &\text{subject to } \Delta P_{k\text{-}l} = \Phi c_{k\text{-}l}. \end{aligned} \tag{14}$$

The general procedure of the proposed ISF estimation method described above is outlined in Fig. 3.

Remark 1: In general, the ISFs of another line $L_{k'\text{-}l'}$, $\Psi_{k'\text{-}l'}$, can be computed from the same set of power injection data as used to compute $\Psi_{k\text{-}l}$. We need to replace $\Delta P_{k\text{-}l}$ in (13) with $\Delta P_{k'\text{-}l'}$, corresponding to the new line of interest. At the same time, we reorder the columns of ΔP according to the electrical distance away from line $L_{k'\text{-}l'}$ to compute the new measurement matrix Φ associated with line $L_{k'\text{-}l'}$. This procedure can be extended to any set of lines of interest. In other words, we can determine ISFs of multiple lines in parallel, by reusing the power injection data ΔP with measurements of active power flow across each line of interest. □

A sufficient condition for a stable solution to (14) for κ-sparse (where κ is the exact, or maximum number of nonzero entries) $c_{k\text{-}l}$ is that the matrix ΔP_s satisfies the *restricted isometry property* (RIP) [15]. The RIP can be achieved with high probability by selecting ΔP_s as a random matrix, the elements of which are independent and identically distributed (i.i.d.) random variables drawn from a Gaussian distribution, if $m \geq \alpha\kappa\log(n/\kappa)$, for some positive $\alpha < 1$ [13], [18]. In our setting, the matrix ΔP_s is constructed from active power injection measurements at all (or a subset of all) buses, and therefore cannot be designed beforehand. However, small variations in the active power injections at each bus can be attributed to random fluctuations in electricity consumption by end users. Hence, we model the measurements of active power injection at each bus with a Gaussian probability density function, as described in Example 3. Moreover, since we assume the power injection fluctuations at each bus are independent, the entries in ΔP (and hence ΔP_s) are, indeed, i.i.d. random variables drawn from a Gaussian distribution. Furthermore, the measurement matrix $\Theta = \Delta P_s \Lambda$ is i.i.d. Gaussian; thus, it has the RIP with high probability regardless of the choice of orthonormal basis Λ [18]. An example of such an orthonormal basis is the wavelet basis, used commonly in image compression applications. Note that even though the transformation M^{-1} is only a basis, not an orthonormal one, we find the

sparse vector recovery results to be satisfactory as evidenced through system case studies.

The problem in (14) is NP-hard (nondeterministic polynomial time-hard) due to the unavoidable combinatorial search [13]. There are numerous classes of computational techniques for solving sparse approximation problems; two major ones are greedy pursuit methods and convex relaxation methods [20], which we describe in the next section as they apply to the solution of (14).

IV. Computation of Post-Transformation ISF Vector via Sparse Methods

In our setting, the signal of interest, the ISF vector $\Psi_{k\text{-}l}$, is not necessarily sparse. Therefore, as described in Section III, we find an appropriate linear coordinate transformation M that results in $c_{k\text{-}l}$, which is approximately sparse. With this approximate sparsity assumption, we solve (14) via greedy pursuit and convex relaxation algorithms.

Greedy pursuit algorithms refine a solution iteratively by choosing one or more components of ΔP_s that yield the greatest improvement until a convergence criterion is met. Convex relaxation algorithms replace the combinatorial problem in (14) with a convex optimization problem by relying on the close approximation of the l_1-norm to the l_0-norm when the vector is sparse. In this section, we consider one specific algorithm from each of these two major classes, namely the orthogonal matching pursuit (OMP) in the class of greedy pursuit algorithms, and a log-barrier algorithm in the class of convex relaxation algorithms.

A. Orthogonal Matching Pursuit

The central idea behind OMP is to successively build a set of the most likely locations of the nonzero terms in the desired vector, $c_{k\text{-}l}$, and then estimate the values of these nonzero entries. OMP is attractive owing to its algorithmic simplicity and provably good approximation accuracy [21]. In this section we describe this algorithm as it applies to finding sparse solutions for (14); its pseudocode is provided in Algorithm 1.

In the p^{th} iteration of OMP, column n_p of Φ, which is most correlated with the approximation error (or residual) at the current step, is identified and added to the set of columns, denoted by Ω_{p-1}, that contains column indices that have already been chosen and added in steps 1 to $p-1$. Since the algorithm chooses column n_p by comparing the correlation between all columns of Φ with the current residual, we normalize the columns of Φ so that each column has unit norm. Denoting this new normalized matrix by $\tilde{\Phi}$, each of its columns can be expressed as

$$\tilde{\varphi}_i = \frac{\varphi_i}{||\varphi_i||_2}, \tag{15}$$

where φ_i is the i^{th} column of Φ. To satisfy the constraint in (13), let $\tilde{c}^i_{k\text{-}l} = c^i_{k\text{-}l}||\varphi_i||_2$, for each $i = 1, \ldots, n$. Corresponding to the columns selected via correlation, the indices in Ω_p also represent the locations of the nonzero entries in $\tilde{c}_{k\text{-}l}$. The values of these nonzero entries are estimated via LSE using only the columns of $\tilde{\Phi}$ present in Ω_p. The orthogonality in this algorithm's namesake manifests itself in the LSE step in that the chosen columns in Ω_p are orthogonal to the residual r_p, and hence these columns will never be chosen again in successive iterations. Finally, the residual is updated based on the new estimation for $\tilde{c}_{k\text{-}l}$ as follows:

$$r_p = \Delta P_{k\text{-}l} - \tilde{\Phi} x_p,$$

where x_p denotes the p^{th} estimate for $\tilde{c}_{k\text{-}l}$, which contains p nonzero entries. This procedure described above is repeated until some stopping criterion.

Algorithm 1

Input: $\Delta P_{k\text{-}l} \in \mathbb{R}^m$, $\tilde{\Phi} \in \mathbb{R}^{m \times n}$.
Output: A sparse vector $x \in \mathbb{R}^n$

1: **Initialize.** Set residual $r_0 = \Delta P_{k\text{-}l}$, index set $\Omega_0 = \emptyset$, and counter $p = 1$
2: **while** $||r_p||_2 > \epsilon$ **do**
3: **Identify.** Find column n_p of $\tilde{\Phi}$ that is most strongly correlated with the residual r_{p-1}:
$n_p = \arg\max_i |\tilde{\varphi}_i^T r_{p-1}|$,
and set $\Omega_p = \Omega_{p-1} \cup n_p$.
4: **Estimate.** Find the best values with the columns chosen so far:
$x_p = \arg\min_y ||\Delta P_{k\text{-}l} - \tilde{\Phi}_{\Omega_p} y||_2$.
5: **Update.** $r_p = \Delta P_{k\text{-}l} - \tilde{\Phi} x_p$
6: **Set.** $p \leftarrow p + 1$
7: **end while**

There are several natural choices for the stopping criterion (see, e.g., [21]). Here we describe two that are relevant to the problem addressed in this paper. First, if the sparsity of the vector $c_{k\text{-}l}$ is unknown, i.e., the number of nonzero entries in the unknown vector is not known a priori, we can terminate the iterations when a sufficiently small residual magnitude $||r_p||_2$ has been reached, or when $p = m$, whichever occurs sooner. On the other hand, if the exact sparsity, or an upper bound to the sparsity of $c_{k\text{-}l}$ is known a priori, we may set the algorithm to halt after a fixed number of iterations $p = \kappa$. In Algorithm 1, we assume no sparsity information is known a priori and set the stopping criterion to depend on the magnitude of the approximation residual.

An additional variation: Suppose, prior to online estimation using OMP, we had conducted either offline model-based studies or online LSE-based computations to obtain base-case ISFs and applied the difference transformation described in (11) to them. We may be privy to not only the sparsity level of the transformed ISF vector representation, but also the locations of the nonzero entries. In this case, we may reach an estimate for $\tilde{c}_{k\text{-}l}$ in one iteration in Algorithm 1.

Finally, let $\hat{\tilde{c}}_{k\text{-}l}$ represent the estimate for $\tilde{c}_{k\text{-}l}$. We recover the estimate for $c_{k\text{-}l}$, denoted by $\hat{c}_{k\text{-}l}$, as

$$\hat{c}^i_{k\text{-}l} = \frac{\hat{\tilde{c}}^i_{k\text{-}l}}{||\varphi_i||_2}, \text{ for each } i = 1, \ldots, n. \tag{16}$$

B. A Log-Barrier Algorithm

Another approach to solve the sparse approximation problem in (14) is to replace the l_0-norm with the l_1-norm, yielding the following convex optimization problem:

$$\begin{aligned} &\min_{c_{k\text{-}l}} ||c_{k\text{-}l}||_1 \\ &\text{subject to } \Delta P_{k\text{-}l} = \Phi c_{k\text{-}l}, \end{aligned} \tag{17}$$

where $||c_{k\text{-}l}||_1 := \sum_{i=1}^{n} |c_{k\text{-}l}^i|$; a problem known as *basis pursuit*, which admits tractable algorithmic solutions (see, e.g., [18], [22]). This program is very effective in recovering signals that are only approximately sparse [23]. While there exists more computationally competitive gradient-based algorithms to solve (17) with very sparse solutions (see e.g., [24]), interior-point methods are used due to their insensitivity to the solution sparsity and robustness against cases of very slow performance or outright failure, which are prevalent in their gradient-based counterparts [20].

To further account for measurement errors that may arise from equipment non-idealities, we relax the equality constraint in (17) to a quadratic constraint. The modified optimization problem is as follows:

$$\begin{aligned} &\min_{c_{k\text{-}l}} ||c_{k\text{-}l}||_1 \\ &\text{subject to } ||\Delta P_{k\text{-}l} - \Phi c_{k\text{-}l}||_2 \leq \epsilon, \end{aligned} \tag{18}$$

where ϵ is a user-specified parameter that may depend on the level of measurement noise expected from the PMU data. It has been shown that for a sufficiently sparse vector $c_{k\text{-}l}$, with $\Delta P_{k\text{-}l} = \Phi c_{k\text{-}l} + e$, for some small error term $||e||_2 \leq \epsilon$, the solution $\hat{c}_{k\text{-}l}$ to (18) is close to $c_{k\text{-}l}$, i.e., $||c_{k\text{-}l} - \hat{c}_{k\text{-}l}||_2 < C\epsilon$, where C is a small constant [15]. Further, to obtain a differentiable objective function, we re-cast the problem in (18) to the following second-order cone program (SOCP) (see, e.g., [25]):

$$\begin{aligned} &\min_{c_{k\text{-}l}, u} \sum_{i=1}^{n} u^i \\ &\text{s.t. } f_i(c_{k\text{-}l}, u) = c_{k\text{-}l}^i - u^i \leq 0,\ i = 1, \ldots, n, \\ &\quad f_{n+i}(c_{k\text{-}l}, u) = -c_{k\text{-}l}^i - u^i \leq 0,\ i = 1, \ldots, n, \\ &\quad f_{2n+1}(c_{k\text{-}l}, u) = \frac{1}{2}\left(||\Delta P_{k\text{-}l} - \Phi c_{k\text{-}l}||_2^2 - \epsilon^2\right) \leq 0, \end{aligned} \tag{19}$$

where $u = [u^1, \ldots, u^n]^T$, and $c_{k\text{-}l} = [c_{k\text{-}l}^1, \ldots, c_{k\text{-}l}^n]^T$ are the decision variables. In order to solve (19), we follow the log-barrier algorithm described in [26], and summarized below; the pseudocode for this algorithm is provided in Algorithm 2. This algorithm expects an initial guess that lies within the feasible region, as delineated by the constraints in (19). In our simulations presented in the remainder of this paper, we choose the starting point for $c_{k\text{-}l}$ to be the one that minimizes $||c_{k\text{-}l}||_2$, namely, $(c_{k\text{-}l})_0 = \Phi^T(\Phi\Phi^T)^{-1}\Delta P_{k\text{-}l}$. The initial point for each u^i, $i = 1, \ldots, n$, must satisfy $-u^i \leq c^i \leq u^i$. The most straightforward starting point may be to set $u_0^i = \max_i |(c_{k\text{-}l})_0^i|$, for all $i = 1, \ldots, n$. In our simulations, the initial condition for each u^i is obtained as follows:

$$u^i = 0.95|c_{k\text{-}l}^i| + 0.1 \max_i |c_{k\text{-}l}^i|.$$

The main idea behind the log-barrier method is to transform (19) into an unconstrained optimization program by incorporating the inequality constraints into the objective function via a penalty function. The ideal penalty function assumes the value 0 if the constraint is satisfied and ∞ otherwise. Since such a function is discontinuous and nondifferentiable, we approximate it using a logarithm relation and transform (19) into the following unconstrained optimization problem:

$$\min_z f(z) = \sum_{i=1}^{n} u_i + \frac{1}{\tau_p} \sum_{i=1}^{2n+1} -\log\left(-f_i(z)\right), \tag{20}$$

where $z = [c_{k\text{-}l}^T, u^T]^T$, $\tau_p > \tau_{p-1}$, and p denotes the log-barrier algorithm iteration index (note that the logarithmic function approaches the ideal penalty function as τ_p increases).

Let $z = z_p + \Delta z$. At each iteration p of the log-barrier algorithm, we form a quadratic approximation of the objective function in (20) as follows:

$$f(z_p + \Delta z) \approx f(z_p) + f_z\big|_{z_p} \Delta z + \Delta z^T H_z\big|_{z_p} \Delta z, \tag{21}$$

where

$$f_z = \begin{bmatrix} 0 & \cdots & 0 & 1 & \cdots & 1 \end{bmatrix}^T - \frac{1}{\tau_p} \sum_{i=1}^{2n+1} \frac{1}{f_i(z)} \nabla f_i(z)$$

and

$$H_z = \frac{1}{\tau_p} \sum_{i=1}^{2n+1} \left[\frac{1}{f_i^2(z)} \nabla f_i(z) \left(\nabla f_i(z)\right)^T - \frac{1}{f_i(z)} \nabla^2 f_i(z) \right].$$

The quadratic approximation in (21) is minimized when

$$f_z\big|_{z_p} + H_z\big|_{z_p} \Delta z = 0, \tag{22}$$

a system of equations that can be solved via Newton's method for Δz. In each log-barrier iteration p, the Newton inner-loop is initialized with z_{p-1} from the previous log-barrier step and so the solution to (22) can be acquired with just a few iterations of Newton's method. If m and n are large, it may be difficult or infeasible to form the matrix H_z and then solve the linear system in (22). However, since (22) is symmetric and positive definite, we can always solve it using a matrix-free method such as the conjugate gradient method [25].

It can be shown that the duality gap associated with $z^*(t)$ and the dual $\lambda^*(t)$ is $(2n+1)/\tau$ (see e.g., [26], Ch. 11). As a consequence, we have, at the p^{th} iteration of the log-barrier algorithm, that

$$\sum_{i=1}^{n} u_p^i - \sum_{i=1}^{n} u^{i*} \leq \frac{2n+1}{\tau_p}. \tag{23}$$

Therefore, the natural stopping criterion of this algorithm is to check whether or not the duality gap has reached a predefined small quantity, η. We note, further, that if τ_p is an increasing sequence, then $z_p \to z^*$ as $\tau_p \to \infty$.

Example 3 (IEEE 14-Bus System): Here, we consider the same system as in Example 1, and use the two algorithms described above to compute estimates for $c_{2\text{-}3}$, in conjunction with simulated PMU data.

For the remainder of this paper, in order to simulate PMU measurements of slight fluctuations in active power injection

Algorithm 2

Input: $\Delta P_{k\text{-}l} \in \mathbb{R}^m$, $\Phi \in \mathbb{R}^{m\times n}$.
Output: A sparse vector $z^* \in \mathbb{R}^n$

1: **Initialize.** Set z_0 within feasible region, τ_1, tolerance η, multiplier μ, and counter $p = 1$
2: **while** $(2n+1)/\tau_p > \eta$ **do**
3: **Newton Step.** Solve (20) with z_{p-1} as initial guess and compute $z_p = z_{p-1} + \Delta z$
4: **Update.** $z^* = z_p$ and $\tau_{p+1} = \mu\tau_p$
5: **Set.** $p \leftarrow p + 1$
6: **end while**

at each bus, we synthesize power injection times-series data as follows; the injection at bus i, denoted by P_i, is

$$P_i[j] = P_i^0[j] + \sigma_1 P_i^0[j] v_1 + \sigma_2 v_2, \tag{24}$$

where $P_i^0[j]$ is the nominal power injection at bus i at instant j, and v_1 and v_2 are pseudorandom values drawn from standard normal distributions with 0-mean and standard deviations $\sigma_1 = 0.1$ and $\sigma_2 = 0.1$, respectively. The first component of variation, $\sigma_1 P_i^0[j] v_1$, represents the inherent fluctuations in generation and load, while the second component, $\sigma_2 v_2$, represents random measurement noise. For each set of bus injection data, we compute the power flow, with the slack bus absorbing all power imbalances, and the active power flow through each line for that particular time. For this example, we produced 21 sets of active power injection and flow data to obtain 20 sets of power fluctuation data and used the first m sets for each case described below.

We compare the results obtained with Algorithm 1 (for $\epsilon = 10^{-5}$), and Algorithm 2 (for $\eta = 10^{-5}$), to those obtained with conventional LSE estimation (see Table I). The model-based benchmark transformed ISFs, $c_{2\text{-}3}$, are listed in column 4 (and also plotted in Fig. 2b). Estimates of $c_{2\text{-}3}$, denoted by $\hat{c}_{2\text{-}3}$, are computed with $m < n$ sets of measurements, and are listed in columns 5–10. Finally, we sparsify the LSE-based ISF vector estimate that results from (6) by using 14 and 20 sets of measurements, and list the elements of the transformed vectors in columns 11 and 12 in Table I, for $m = 14, 20$, respectively.

By inspecting Table I, we conclude that Algorithm 2 achieves higher accuracy than Algorithm 1, which is verified by comparing the l_2-norm of the error in each estimate as compared to the benchmark model-based $c_{2\text{-}3}$, as shown in the last row. Further, Algorithm 2 with $m = 10, 13$ leads to more accurate estimates than those obtained using the LSE-based approach with $m = 14$. On the other hand, with Algorithm 1, only the estimate obtained with $m = 13$ is superior to the LSE-based approach with $m = 14$. However, we note that the estimate obtained using the LSE-based method with $m = 20$ is more accurate than all of the estimates obtained with $m < n$. Thus, there is some tradeoff between the level of accuracy and the number of measurements required. In this example, the reduction in the number of required measurements to achieve accuracy level comparable to the LSE-based method is not significant. However, as we show in Section V, via large-scale test cases, this reduction becomes more notable as the size of the system grows.

TABLE II: Comparison between $\hat{\Psi}_{k\text{-}l}$ and $\Psi_{k\text{-}l}$ for several lines in the IEEE 300-bus system.

Line	$\|\|\hat{\Psi}_{k\text{-}l} - \Psi_{k\text{-}l}\|\|_2$			
$L_{272\text{-}268}$	$m = 110$	$m = 120$	$m = 130$	$m = 140$
via Algorithm 1	0.0676	0.0693	0.0551	0.0777
via Algorithm 2	0.0320	0.0341	0.0271	0.0279
via LSE	$m = 305$	$m = 350$	$m = 400$	$m = 600$
	0.5891	0.0962	0.0678	0.0339
$L_{30\text{-}73}$	$m = 190$	$m = 210$	$m = 230$	$m = 250$
via Algorithm 1	0.0640	0.0649	0.0517	0.0494
via Algorithm 2	0.0320	0.0323	0.0268	0.0288
via LSE	$m = 305$	$m = 350$	$m = 400$	$m = 600$
	0.2719	0.0372	0.0236	0.0162
$L_{112\text{-}148}$	$m = 230$	$m = 250$	$m = 270$	$m = 290$
via Algorithm 1	0.9216	0.8255	0.7809	0.0964
via Algorithm 2	0.0849	0.0716	0.0377	0.0474
via LSE	$m = 305$	$m = 350$	$m = 400$	$m = 600$
	0.1414	0.0289	0.0168	0.0116

While ISF estimation results for other lines are not shown, they can be computed from the same set of power injection data as was used to compute $\hat{\Psi}_{2\text{-}3}$ above. We only need to replace $\Delta P_{k\text{-}l}$ in (13) with the vector that results from flow measurements that correspond to the new line of interest. At the same time, we reorder columns of ΔP according to the electrical distance away from the new line of interest. ■

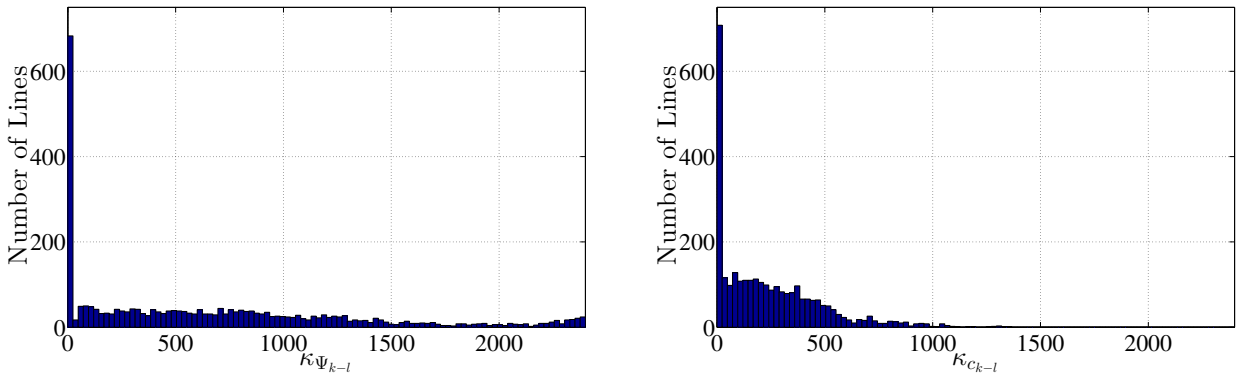
V. Case Studies

In this section, we further illustrate the concepts presented in Sections III and IV using the IEEE 300-bus test system and the 2383-bus Polish power system. The simulation tool MATPOWER [27] is used throughout to compute relevant transmission line flow measurements from some synthetic power injection profiles generated using (24).

A. IEEE 300-Bus System

For this case study, we focus on the ISFs of lines $L_{272\text{-}268}$, $L_{30\text{-}73}$, and $L_{112\text{-}148}$ with respect to each bus in the system. These lines are selected because they are typical of ISF vectors in the system corresponding to three different levels of $c_{k\text{-}l}$ sparsity, with $c_{272\text{-}268}$ most sparse and $c_{112\text{-}148}$ least. Using a model of the system, we first compute benchmark values for $\Psi_{272\text{-}268}$, $\Psi_{30\text{-}73}$, and $\Psi_{112\text{-}148}$ based on the definition of the ISF in (2), to which we compare estimates obtained via greedy pursuit, convex relaxation, and LSE algorithms. These original ISF vectors, ordered by the somewhat arbitrary network diagram designation (see, e.g., [16]), is neither sparse nor particularly smooth. As in Example 3, we simulate PMU measurements of slight fluctuations in active power injection at each bus via (24), from which we obtain the matrix ΔP, the regressor matrix used in ISF estimation for any line of interest. Next, we reorder the elements of each ISF vector by the electrical distance of each node to the line of interest and subsequently the columns of ΔP to obtain ΔP_s. Furthermore, we compute the power flowing through each of lines $L_{272\text{-}268}$, $L_{30\text{-}73}$, and $L_{112\text{-}148}$ for each set of bus injection data. From these computations, we obtain observation vectors $\Delta P_{272\text{-}268}$, $\Delta P_{30\text{-}73}$, and $\Delta P_{112\text{-}148}$, which are then substituted into (13) to solve via Algorithms 1 and 2.

To assess the effectiveness of our proposed sparse vector recovery methods, we choose several values of $m < n$, corresponding to the number of synchronized data sets available.



(a) Sparsity of pre-transformation $\Psi_{k\text{-}l}$ vectors.

(b) Sparsity of post-transformation $c_{k\text{-}l}$ vectors.

Fig. 4: Polish 2383-Bus System: Sparsity of pre- and post-transformation ISFs.

For each test case, we select m rows from ΔP, reorder columns of ΔP by electrical distance away from line $L_{k\text{-}l}$, and finally multiply by the inverse difference transformation matrix to obtain the measurement matrix Φ, as in (13). Combined with active line flows computed via MATPOWER, we solve the sparse vector recovery problem in (13) using the two algorithms described in Section IV: Algorithm 1 in the class of greedy pursuit algorithms (with $\epsilon = 10^{-5}$), and Algorithm 2 in the class of convex optimization algorithms (with $\eta = 10^{-5}$). The resulting vectors are sparse representations, denoted by $\hat{c}_{k\text{-}l}$, of the ISF vector estimates, denoted by $\hat{\Psi}_{k\text{-}l,s}$, which we obtain by applying the inverse difference transformation to $\hat{c}_{k\text{-}l}$, as described in (12). For comparison, we also obtain estimates $\hat{\Psi}_{k\text{-}l}$ directly using the LSE-based method described in [8]. Finally, we evaluate the accuracy of these three methods against the benchmark ISF vectors obtained via direct model-based computation using the definition in (2). The results are described below and summarized in Table II.

In the case of line $L_{272\text{-}268}$, we compute estimates via Algorithms 1 and 2 with $m = 110, 120, 130, 140$, and LSE with $m = 305, 350, 400, 600$. Algorithm 2 results in higher accuracy with only 130 sets of synchronized measurements than LSE with 600 sets. On the other hand, in the case of line $L_{30\text{-}73}$, the estimate obtained via Algorithm 2 with 230 sets of measurements is about as accurate as that obtained using LSE with 400 sets. Regrettably, in the case of line $L_{112\text{-}148}$, estimates obtained via the proposed sparsity-exploiting methods with as many as 290 sets of measurements are unable to surpass those obtained via the LSE method with as few as 350 sets. This decline in performance can be attributed to the sparsity level of the post-transformation ISF vector, $c_{k\text{-}l}$. In this case study, $c_{272\text{-}268}$ is the most sparse, followed by $c_{30\text{-}73}$, and then $c_{112\text{-}148}$. For lines for which the post-transformation ISF vector is not sparse, it may be prudent to search for an alternate representation that is sparse.

We also note that, in general, estimates obtained via Algorithm 2 are more accurate than those via Algorithm 1 using the same observation matrix. This stems from the interior point convex optimization's insensitivity to the sparsity of the solution, so it is able to produce good estimates even for non-sparse solutions, given enough measurement sets. As we show in the next case study involving the 2383-bus Polish power system, while Algorithm 2 may be more robust, Algorithm 1 is more superior if the solution is sparse.

B. Polish 2383-Bus System

For this large-scale test case with 2896 lines, we evaluate the effectiveness of the electrical distance-based method proposed in Section III to sparsify ISF vectors as well as the algorithms described in Section IV to estimate resulting sparse vectors.

1) Sparsify via electrical distance-based method: We consider the benchmark ISFs for all lines in the system computed by linearizing the nonlinear AC power flow equations. We obtain a measure for the sparsity of the pre-transformation ISF vector $\Psi_{k\text{-}l}$ by counting the number of nonzero elements in each normalized ISF vector with[3]

$$\kappa_{\Psi_{k\text{-}l}} = \#\left\{ i \in [1,n] : \frac{|\Psi^i_{k\text{-}l}|}{||\Psi_{k\text{-}l}||_\infty} > 0.005 \right\}. \qquad (25)$$

Similarly, we obtain a comparable sparsity measure for the post-transformation vector $c_{k\text{-}l}$ with

$$\kappa_{c_{k\text{-}l}} = \#\left\{ i \in [1,n] : \frac{|c^i_{k\text{-}l}|}{||c_{k\text{-}l}||_\infty} > 0.005 \right\}. \qquad (26)$$

The sparsity measures in (25) and (26) for all lines are visualized as histograms in Fig. 4a and Fig. 4b, respectively. By comparing the aforementioned figures, we note that there is an increase in the number of vectors with fewer nonzero entries in the post-transformation $c_{k\text{-}l}$'s in Fig. 4b as compared to the pre-transformation $\Psi_{k\text{-}l}$'s in Fig. 4a. In fact, $\kappa_{c_{k\text{-}l}} < 600$ for 92% of the lines, whereas $\kappa_{\Psi_{k\text{-}l}} < 600$ for 54%. Also, we note that the pre-transformation ISF vectors of many lines are already highly sparse, as shown by the leftmost bar in Fig. 4a. This is in agreement with the intuition that, in a large-scale system, many line flows are significantly affected by only a small number of nodal injections. In these cases, the transformation proposed in Section III does not affect the sparsity of the corresponding post-transformation ISF vector.

[3] $\#\mathcal{A}$ denotes the cardinality of set $\mathcal{A}$.

TABLE III: Comparison between $\hat{\Psi}_{k\text{-}l}$ and $\Psi_{k\text{-}l}$ for the Polish 2383-bus system case study.

Line	$\|\|\hat{\Psi}_{k\text{-}l} - \Psi_{k\text{-}l}\|\|_2$			
$L_{53\text{-}52}$	$m = 500$	$m = 600$	$m = 700$	$m = 800$
via Algorithm 1	0.0490	0.0480	0.0413	0.0402
via Algorithm 2	0.0562	0.0516	0.0459	0.0391

TABLE IV: Timing comparison for Polish 2383-bus system case study.

Line	Execution Time [s]			
$L_{53\text{-}52}$	$m = 500$	$m = 600$	$m = 700$	$m = 800$
via Algorithm 1	0.2798	0.4502	0.5263	0.7828
via Algorithm 2	15.5302	13.7511	15.8118	16.1992

2) Sparse ISF vector recovery: To illustrate the effectiveness of the proposed estimation algorithms, we consider the ISFs of line $L_{53\text{-}52}$ with respect to each of the 2383 buses in the system. Due to the size of this test system and the resulting exorbitant simulation time, we refrain from computing $\hat{\Psi}_{k\text{-}l}$ estimates directly using the LSE-based method. As in the 300-bus case study, we exploit a sparse representation of the ISF vector and compute estimates via Algorithm 1 (for $\epsilon = 0.01$), and Algorithm 2 (for $\eta = 10^{-5}$). Then we compare these to benchmark ISF obtained using the model-based approach; the results are shown in Table III. Here, unlike in the 300-bus case study, the solution is highly sparse and so Algorithm 1 results in more accurate ISF estimates than Algorithm 2 from the same synchronized measurement sets (except for the case $m = 800$). Also, as shown in Table IV, Algorithm 1 incurs much lower computation time than Algorithm 2. Next, we further take advantage of the sparsity of the solution and offer a modification to the proposed algorithms.

VI. ISF Estimation with a Subset of Measurements

One of the technical restrictions the framework presented thus far is that bus injection measurements are required at all buses in the system. In a practical setting, however, the entire set of PMU measurements may not be wholly available at every sample time. Fortunately, since bus injections at buses that are geographically or electrically distant from the line of interest would likely not have a significant effect on its flow, it is prudent to assume the ISFs of the line $L_{k\text{-}l}$ with respect to a subset of buses are negligible in magnitude. To this end, denote the set of all buses in the system as $\mathcal{B}$. Suppose we are interested in the flow across the line $L_{k\text{-}l}$, where $k, l \in \mathcal{B}_1 \subseteq \mathcal{B}$. With these definitions, we consider the following modification to (5):

$$\Delta P_{k\text{-}l} = \begin{bmatrix} \Delta P_{\mathcal{B}_1} & \Delta P_{\mathcal{B}_1^c} \end{bmatrix} \begin{bmatrix} \Psi_{k\text{-}l}^{\mathcal{B}_1} \\ \Psi_{k\text{-}l}^{\mathcal{B}_1^c} \end{bmatrix},$$

with $\mathcal{B}_1^c := \mathcal{B} \backslash \mathcal{B}_1$ and where $\Delta P_{\mathcal{B}_1}$ and $\Delta P_{\mathcal{B}_1^c}$ represent the active power injection fluctuation at buses in $\mathcal{B}_1$ and $\mathcal{B}_1^c$, respectively. Similarly, $\Psi_{k\text{-}l}^{\mathcal{B}_1}$ and $\Psi_{k\text{-}l}^{\mathcal{B}_1^c}$ represent the ISFs of line $L_{k\text{-}l}$ with respect to buses in $\mathcal{B}_1$ and $\mathcal{B}_1^c$, respectively. Now, suppose $\mathcal{B}_1$ contains buses such that entries of $\Psi_{k\text{-}l}^{\mathcal{B}_1^c}$ are negligibly small compared to those of $\Psi_{k\text{-}l}^{\mathcal{B}_1}$, then we obtain

$$\Delta P_{k\text{-}l} = \Delta P_{\mathcal{B}_1} \Psi_{k\text{-}l}^{\mathcal{B}_1} + \theta, \tag{27}$$

where $\theta = \Delta P_{\mathcal{B}_1^c} \Psi_{k\text{-}l}^{\mathcal{B}_1^c}$ can be viewed as a measurement error.

TABLE V: Comparison between $\hat{\Psi}_{k\text{-}l}$ and $\Psi_{k\text{-}l}$ for the Polish 2383-bus system case study using a subset of data.

Line	$\|\|\hat{\Psi}_{k\text{-}l} - \Psi_{k\text{-}l}\|\|_2$			
$L_{53\text{-}52}$	$m = 500$	$m = 600$	$m = 700$	$m = 800$
via Algorithm 1	0.0509	0.0482	0.0440	0.0420
via Algorithm 2	0.0579	0.0525	0.0565	0.0632

TABLE VI: Timing comparison for Polish 2383-bus system case study using a subset of data.

Line	Execution Time [s]			
$L_{53\text{-}52}$	$m = 500$	$m = 600$	$m = 700$	$m = 800$
via Algorithm 1	0.2547	0.4706	0.6899	0.9705
via Algorithm 2	18.6132	15.6137	17.2752	19.7156

In order to select the number of buses to include in $\mathcal{B}_1$, i.e., $n_1 = \#\mathcal{B}_1$, we consider the post-transformation ISF vector $c_{k\text{-}l}$ obtained from the model-based sensitivities $\Psi_{k\text{-}l}$. Specifically, n_1 can be chosen such that

$$\frac{|c_{k\text{-}l}^i|}{||c_{k\text{-}l}||_\infty} < \gamma, \ \forall \ i > n_1,$$

for some user-defined $\gamma > 0$. Even though this modification reduces the number of measurements required, PMU measurements must be available at all n_1 buses in $\mathcal{B}_1$.

Example 4 (Polish 2383-Bus System): To illustrate ideas presented above, we consider ISFs of line $L_{53\text{-}52}$ in the Polish power system again. Here, $n_1 = 1000$ buses that are electrically nearest to line $L_{53\text{-}52}$ are chosen to be in $\mathcal{B}_1$, corresponding to $\gamma \approx 0.005$. As shown in Tables V and VI, the performance of this reduced case is similar to the case with full knowledge of all bus injection data, both in accuracy as well as computation time. There are just as many nonzero elements in the reduced case as before; hence, we would not expect a significant reduction in computation time. Moreover, PMU measurements are required at all n_1 buses in $\mathcal{B}_1$. However, the advantage of using a subset of bus injection measurements lies in the reduction in the number of measurements required and consequently communication overhead incurred. ■

VII. Concluding Remarks

In this paper, we presented a method to estimate a vector of ISFs by exploiting a sparse representation of it, and solving for the sparse vector via greedy pursuit and convex optimization. An advantage of the proposed method is that it does not rely on a power flow model of the system, but instead only uses PMU measurements collected in real-time. A direct consequence is that the proposed method is also independent of slack bus location designation. Apart from eliminating the power flow model, we show that the proposed measurement-based approach provides accurate estimates of the ISFs using fewer measurements than those obtained using LSE.

We illustrate the application of this method to the IEEE 14- and 300-bus system power flow models, from which comparisons are made between results obtained via greedy pursuit, convex optimization, and LSE. We also demonstrate the scalability of the proposed method via the Polish 2383-bus system model.

As future work, we plan to conduct more rigorous analysis of the exact or a lower bound to the number of measurement sets required in the proposed underdetermined solution framework. Additionally, future work should investigate other transformations to sparsify the ISF vector for cases in which the proposed electrical distance-based method is inadequate. Moreover, we may be able to apply the estimated ISFs for detecting model-related errors and for obtaining reduced system equivalents. Also of interest are modifications to our proposed method to accommodate measurement configurations based on existing optimal PMU placement methods.

REFERENCES

[1] FERC and NERC. (2012, Apr.) Arizona-southern california outages on september 8, 2011: Causes and recommendations. [Online]. Available: http://www.ferc.gov/legal/staff-reports/04-27-2012-ferc-nerc-report.pdf

[2] U.S.-Canada Power System Outage Task Force. (2004, Apr.) Final report on the august 14th blackout in the united states and canada: causes and recommendations. [Online]. Available: https://reports.energy.gov/BlackoutFinal-Web.pdf

[3] B. Stott, J. Jardim, and O. Alsac, "Dc power flow revisited," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1290 – 1300, 2009.

[4] A. Wood and B. Wollenberg, *Power Generation, Operation and Control*. New York: Wiley, 1996.

[5] J. E. Tate and T. J. Overbye, "Line outage detection using phasor angle measurements," *IEEE Transactions on Power Systems*, vol. 23, no. 4, pp. 1644 – 1652, 2008.

[6] H. Zhu and G. B. Giannakis, "Sparse overcomplete representations for efficient identification of power line outages," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2215 – 2224, 2012.

[7] R. Emami and A. Abur, "External system line outage identification using phasor measurement units," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1035 – 1040, 2013.

[8] Y. C. Chen, A. D. Domínguez-García, and P. W. Sauer, "Online computation of power system linear sensitivity distribution factors," in *Proc. of the IREP Bulk Power System Dynamics and Control Symposium*, Aug. 2013.

[9] ——, "Measurement-based estimation of linear sensitivity distribution factors and applications," *IEEE Transactions on Power Systems*, vol. 29, no. 3, pp. 1372 – 1382, 2014.

[10] US DOE Electricity Delivery & Energy Reliability. (2013, Aug.) Synchrophasor technologies and their deployment in the recovery act smart grid programs. [Online]. Available: http://energy.gov/sites/prod/files/2013/08/f2/SynchrophasorRptAug2013.pdf

[11] E. Schweitzer, D. Whitehead, and G. Zweigle, "Practical synchronized phasor solutions," in *Proc. of the IEEE Power Energy Society General Meeting*, July 2009, pp. 1–8.

[12] Y. C. Chen, A. D. Domínguez-García, and P. W. Sauer, "Online estimation of power system distribution factors—a sparse representation approach," in *Proc. of the North American Power Symposium*, Sep. 2013.

[13] Y. C. Eldar and G. Kutyniok, *Compressed sensing: theory and applications*. Cambridge University Press, 2012.

[14] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289 – 1306, 2006.

[15] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489 – 509, 2006.

[16] University of Washington. (2012, Oct.) Power system test case archive. [Online]. Available: http://www.ee.washington.edu/research/pstca/

[17] A. Conejo, J. Contreras, D. Lima, and A. Padilha-Feltrin, "Zbus transmission network cost allocation," *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 342 – 349, Feb. 2007.

[18] R. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118 – 121, Jul. 2007.

[19] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*. Marcel Dekker, 2004.

[20] J. Tropp and S. Wright, "Computational methods for sparse solution of linear inverse problems," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 948 – 958, Jun. 2010.

[21] J. Tropp, "Greed is good: algorithmic results for sparse approximation," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2231 – 2242, Oct. 2004.

[22] D. Needell and R. Vershynin, "Greedy signal recovery and uncertainty principles," in *Proc. of SPIE - The International Society for Optical Engineering*, vol. 6814, 2008.

[23] J. Romberg, "Imaging via compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 14 – 20, Mar. 2008.

[24] S. Sardy, A. Bruce, and P. Tseng, "Block coordinate relaxation methods for nonparametric wavelet denoising," *Journal of Computational and Graphical Statistics*, vol. 9, no. 2, pp. 361 – 379, 2000.

[25] E. J. Candès and J. Romberg, "l_1-MAGIC: Recovery of sparse signals via convex programming," California Institute of Technology, Tech. Rep., Feb. 2005.

[26] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, 2004.

[27] R. D. Zimmerman, C. E. Murillo-Snchez, and R. J. Thomas, "Matpower: 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.

Yu Christine Chen (S'10) received the B.A.Sc. degree in Engineering Science (major in Electrical Engineering) from the University of Toronto in Canada in 2009 and the M.S. degree in Electrical Engineering from the University of Illinois at Urbana-Champaign in 2011. She is currently pursuing a Ph.D. degree in Electrical Engineering at the University of Illinois at Urbana-Champaign.

Her research interests include power system dynamics and monitoring, and renewable resource integration.

Alejandro D. Domínguez-García (S'02, M'07) received the degree of Electrical Engineer from the University of Oviedo (Spain) in 2001 and the Ph.D. degree in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, MA, in 2007.

He is an Associate Professor in the Electrical and Computer Engineering Department at the University of Illinois at Urbana-Champaign, where he is affiliated with the Power and Energy Systems area; he also has been a Grainger Associate since August 2011. He is also an Associate Professor in the Coordinated Science Laboratory and in the Information Trust Institute, both at the University of Illinois at Urbana-Champaign. His research interests are in the areas of system reliability theory and control, and their applications to electric power systems, power electronics, and embedded electronic systems for safety-critical/fault-tolerant aircraft, aerospace, and automotive applications.

Dr. Domínguez-García received the NSF CAREER Award in 2010, and the Young Engineer Award from the IEEE Power and Energy Society in 2012. He is an editor of the IEEE TRANSACTIONS ON POWER SYSTEMS and the IEEE POWER ENGINEERING LETTERS.

Peter W. Sauer (S'73, M'77, SM'82, F'93, LF'12) received the B.S. degree from the University of Missouri at Rolla, Rolla, MO, USA, in 1969, and the M.S. and Ph.D. degrees from Purdue University, West Lafayette, IN, USA, in 1974 and 1977, respectively, all in electrical engineering.

From 1969 to 1973, he was an Electrical Engineer on a design assistance team for the Tactical Air Command at Langley Air Force Base, VA, USA, working on design and construction of airfield lighting and electrical distribution systems. He has been on the faculty at The University of Illinois at Urbana-Champaign, Urbana, IL, USA, since 1977, where he teaches courses and directs research on power systems and electric machines. From August 1991 to August 1992, he served as the Program Director for Power Systems in the Electrical and Communication Systems Division, National Science Foundation, Washington, DC, USA. He is a cofounder of the Power Systems Engineering Research Center (PSERC) and has served as the Illinois site director from 1996 to the present. He is a cofounder of PowerWorld Corporation and served as Chairman of the Board of Directors from 1996 to 2001. He is currently the Grainger Chair Professor of Electrical Engineering at Illinois.

Dr. Sauer is a registered Professional Engineer in Virginia and Illinois and a member of the U.S. National Academy of Engineering.