

Article

State Estimation in Partially Observable Power Systems via Graph Signal Processing Tools

Lital Dabush ^{1,*} , Ariel Kroizer ¹ and Tirza Routtenberg ^{1,2} 

¹ School of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer Sheva 8410501, Israel

² Department of Electrical and Computer Engineering, Princeton University, Princeton, NJ 08544, USA

* Correspondence: litaldab@post.bgu.ac.il

Abstract: This paper considers the problem of estimating the states in an unobservable power system, where the number of measurements is not sufficiently large for conventional state estimation. Existing methods are either based on pseudo-data that is inaccurate or depends on a large amount of data that is unavailable in current systems. This study proposes novel graph signal processing (GSP) methods to overcome the lack of information. To this end, first, the graph smoothness property of the states (i.e., voltages) is validated through empirical and theoretical analysis. Then, the regularized GSP weighted least squares (GSP-WLS) state estimator is developed by utilizing the state smoothness. In addition, a sensor placement strategy that aims to optimize the estimation performance of the GSP-WLS estimator is proposed. Simulation results on the IEEE 118-bus system show that the GSP methods reduce the estimation error magnitude by up to two orders of magnitude compared to existing methods, using only 70 sampled buses, and increase of up to 30% in the probability of bad data detection for the same probability of false alarms in unobservable systems. The results conclude that the proposed methods enable an accurate state estimation, even when the system is unobservable, and significantly reduce the required measurement sensors.

Keywords: graph signal processing (GSP); power system state estimation (PSSE); network observability; sensor allocation



Citation: Dabush, L.; Kroizer, A.; Routtenberg, T. State Estimation in Partially Observable Power Systems via Graph Signal Processing Tools. *Sensors* **2023**, *23*, 1387. <https://doi.org/10.3390/s23031387>

Academic Editor: Toan Phung

Received: 18 December 2022

Revised: 8 January 2023

Accepted: 22 January 2023

Published: 26 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Power system state estimation (PSSE) is a critical component of modern energy management systems (EMSs) for multiple purposes, including monitoring, analysis, security, control, and management of the power delivery [1]. The PSSE is conducted using topological information, power measurements, and physical constraints to estimate the voltages (states) at the system buses. The performance and reliability of the PSSE largely depend on the availability and the quality of the measurements [2]. However, there are various realistic scenarios where the system is *partially observable* (also named in the literature as unobservable) ([1] Chapter 4, [3–5]), that is, the number of sensors is not sufficiently large, or sensors are not well placed in the network. The observability of the system may be compromised due to communication errors, topology changes, sensor failures [1], malicious attacks [6–8], and electrical blackouts [9]. A direct implication of system unobservability is that conventional estimators that assume deterministic states, such as the commonly used weighted least squares (WLS) estimator, can no longer be used since they are inaccurate, inconsistent, and may have large estimation errors even in the absence of noise [3,10]. Therefore, developing new estimation methods that enable the full functionality of systems without requiring observability is crucial for reliable operation of the power grid.

State estimation in partially observable systems must incorporate additional properties or information beyond the power flow equations in order to obtain a meaningful estimation. Most existing approaches are two-step solutions that first produce additional

pseudo-measurements, e.g., based on short-term load forecasting, to make the system observable, and then estimate the states using an existing technique [11–14]. However, pseudo-measurements do not contain real-time data, and thus result in an inaccurate estimation. Dynamic state estimation utilizes measurements at different time points [14,15], but needs fast scan rates to capture the dynamics, and is also based on restrictive stationary system assumptions [14]. PSSE that uses data from smart meters and phasor measurement units (PMUs) to overcome observability issues [10,16–18] is usually inapplicable, due to the limited deployment of these sensors [19], and their initial installment cost [20,21]. Sparse signal recovery methods [5] use matrix completion to estimate the states under low observability conditions. However, to employ the matrix completion framework, the measurements matrix should be low-rank. Unfortunately, this assumption is system-dependent and does not always hold, due to, e.g., the spatial correlation between loads at neighboring buses. Deep learning techniques have recently been used for pseudo-measurement generation, and for the reconstruction of missing data for PSSE [19,22,23]. However, these techniques heavily depend on the availability of numerous high-quality event labels that are rarely available in practice [24,25]. In addition, some of these methods do not utilize the physical model [19], which may result in poor performance in practice.

Concepts from graph theory and graphical models have been used in power systems for sensor placement [26], topology identification [27–29], state estimation [6,30,31], analysis [32,33], and optimal power flow calculation [34,35]. However, the graphical model methods assume a specific statistical structure, which does not necessarily apply in power systems, and often do not use the physical models, which may result in poor performance in practice. GSP is an emerging field that extends concepts and techniques from traditional digital signal processing (DSP) to data on graphs [36–39]. Recent works have proposed the integration of GSP in power systems, such as the application of the GSP framework for the power grid with PMU data in [40], the spectral graph analysis of power flows in [35], and attack detection by GSP in [7,41,42]. For unobservable power systems, the authors of [43] investigate the use of GSP methods for dynamic state estimation in unobservable systems based on PMUs and advanced metering infrastructures (AMIs) under the assumption of bandlimited graph signals. However, state estimation without full observability that does not depend on PMUs and AMIs using GSP tools has not been demonstrated before. Therefore, using GSP tools has a great potential for overcoming the lack of information in state estimation without full observability.

This research develops new GSP methods for state estimation in power systems based on interpreting the voltage signals (phases and magnitudes) as graph signals. First, it is shown empirically and analytically that the states for static PSSE, i.e., voltages, are smooth graph signals with respect to the nodal admittance matrix, which is a Laplacian matrix in the graph representation of the network. Second, a GSP-WLS estimation method is developed for PSSE in the direct current power flow (DC-PF) model that uses the graph smoothness of the states and does not require the full observability of the network. Next, a new approach for sensor placement is introduced in order to optimize the estimation performance obtained by the GSP-WLS estimator. Finally, the proposed estimation method is extended to the more realistic alternating current power flow (AC-PF) model by developing a regularized Gauss–Newton method for PSSE that uses the smoothness of the voltage phases and magnitudes. The simulations show that the proposed methods can accurately estimate voltage phases and magnitudes, and detect bad data under conditions of low observability, where standard methods cannot (or display poor performance).

The aim of this paper is to establish a GSP framework for state estimation, sensor allocation, and bad data detection in partially observable systems. The key novelties are as follows:

- It is demonstrated that the voltages in the power system can be represented as smooth graph signals, where the graph Laplacian is the admittance matrix. This result can serve as a foundation for developing new GSP tools for other power systems applications in future research.

- New state estimation methods for PSSE in both DC-PF and AC-PF models are developed. These methods use the graph smoothness of the states, and do not require the full observability of the network. While regularization using the Laplacian quadratic form has been applied in various applications, such as image processing [44,45], principal component analysis (PCA) [46], data classification [47,48], and semisupervised learning on graphs [49,50], it has not been conducted before in the context of unobservable power systems. Additionally, the nonlinear measurement equations in the AC-PF model present a new challenge from a GSP perspective, requiring the incorporation of graphical information in the form of Laplacian regularization into the iterative method. As such, these state estimation methods contribute to the expansion of the GSP toolbox for a wide range of applications.
- A new approach for sensor placement is introduced to optimize the estimation performance. As the mean squared error (MSE) of the estimator depends on the unknown state vector, the minimization of the Cramér–Rao bound (CRB) is utilized instead. This results in a novel approach that can potentially be applied to other applications in the future.
- Numerical simulations on the IEEE 118-bus system are used to validate the merit of the new estimators and the new sensor placement method under different setups, compared to existing pseudo-measurement and matrix completion techniques.

The rest of this paper is organized as follows. In Section 2, the GSP background is introduced, as well as the model, and the conventional estimation approach. In Section 3, the GSP properties of the states are studied. In Section 4, the GSP-WLS state estimator is derived for the DC-PF model and a sensor placement method. In Section 5, the proposed estimation method to the AC-PF model is extended by deriving the regularized Gauss–Newton method. A simulation study is presented in Section 6, and the discussion and conclusions are provided in Section 8.

2. Background and Model

A power system can be represented as an undirected weighted graph, where the nodes and the edges of the graph are the grid buses and transmission lines, respectively. This section begins with background on the theory of GSP in Section 2.1. Then, the considered power flow measurement model, as well as the state estimation and network observability for this model, are presented in Section 2.2.

In the rest of this paper, vectors and matrices are denoted by boldface lowercase letters and boldface uppercase letters, respectively. The notations $(\cdot)^T$, $(\cdot)^{-1}$, $(\cdot)^\dagger$, and $\text{Tr}(\cdot)$ denote the transpose, inverse, Moore–Penrose pseudo-inverse, and trace operators, respectively. The m th element of the vector \mathbf{a} and the (m, q) th element of the matrix \mathbf{A} are denoted by a_m and $A_{m,q}$, respectively. Similarly, $\mathbf{A}_{\mathcal{S}_1, \mathcal{S}_2}$ denotes the submatrix of \mathbf{A} , the rows and columns of which are indexed by the sets \mathcal{S}_1 and \mathcal{S}_2 , where $\mathbf{A}_{\mathcal{S}} \triangleq \mathbf{A}_{\mathcal{S}, \mathcal{S}}$, and $\mathbf{a}_{\mathcal{S}}$ is a subvector of \mathbf{a} containing the elements indexed by \mathcal{S} . The cardinality of the set \mathcal{S} is denoted by $|\mathcal{S}|$. The gradient of a vector function $\mathbf{g}(\mathbf{x}) \in \mathbb{R}^K$ with respect to $\mathbf{x} \in \mathbb{R}^M$, $\frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}}$, is a matrix in $\mathbb{R}^{K \times M}$, with the (k, m) th element equal to $\frac{\partial g_k}{\partial x_m}$. The matrices \mathbf{I} and $\mathbf{0}$ denote the identity matrix and the zero matrix with appropriate dimensions, respectively, and $\|\cdot\|$ denotes the Euclidean l_2 -norm.

2.1. Background: GSP Framework

Let $\mathcal{G}(\mathcal{V}, \xi)$ be a general undirected weighted graph, where $\mathcal{V} = \{1, \dots, N\}$ and $\xi = \{1, \dots, P\}$ are the sets of nodes and edges, respectively. The matrix $\mathbf{W} \in \mathbb{R}^{N \times N}$ is the weighted adjacency matrix of the graph $\mathcal{G}(\mathcal{V}, \xi)$, where $W_{k,n}$ denotes the weight of the edge between node k and node n . It is assumed that $W_{k,n} \geq 0$ and that $W_{k,n} = 0$ if no edge exists between k and n . The graph Laplacian matrix is defined as

$$L_{k,l} = \begin{cases} \sum_{n=1}^N W_{k,n}, & k = l \\ -W_{k,l}, & \text{otherwise} \end{cases}, \quad k, l = 1, \dots, N. \quad (1)$$

The Laplacian matrix is a positive semidefinite matrix with the eigenvalue decomposition

$$\mathbf{L} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T, \quad (2)$$

where the columns of \mathbf{V} are the eigenvectors of \mathbf{L} , $\mathbf{V}^T = \mathbf{V}^{-1}$, and $\mathbf{\Lambda}$ is a diagonal matrix consisting of the ordered eigenvalues of \mathbf{L} : $0 = \lambda_1 < \lambda_2 \leq \dots \leq \lambda_N$. By analogy to signal frequency in DSP, the Laplacian eigenvalues can be interpreted as the graph frequencies that, together with the eigenvectors in \mathbf{V} , define the spectrum of the graph $\mathcal{G}(\mathcal{V}, \xi)$ [37].

A graph signal is a function that assigns a scalar value to each node, and thus, is an N -dimensional vector. The graph Fourier transform (GFT) of a graph signal \mathbf{a} with respect to the graph $\mathcal{G}(\mathcal{V}, \xi)$ is [37]

$$\tilde{\mathbf{a}} \triangleq \mathbf{V}^{-1}\mathbf{a}. \quad (3)$$

Similarly, the inverse GFT is obtained by left multiplication of a vector by \mathbf{V} . The Dirichlet energy of a graph signal, \mathbf{a} , is defined as

$$\mathcal{E}_{\mathbf{L}}(\mathbf{a}) \triangleq \mathbf{a}^T \mathbf{L} \mathbf{a} = \frac{1}{2} \sum_{k=1}^N \sum_{n=1}^N W_{k,n} (a_k - a_n)^2 = \sum_{k=1}^N \lambda_k \tilde{a}_k^2, \quad (4)$$

where the second equality is obtained by substituting (1), and the last equality is obtained by substituting (2) and (3). The Dirichlet energy is a smoothness measure, which is used to quantify the variability encoded by the graph weights [37]. A graph signal, \mathbf{a} , is smooth if

$$\mathcal{E}_{\mathbf{L}}(\mathbf{a}) \leq \varepsilon, \quad (5)$$

where ε is small in terms of the specific application [37]. It can be seen that the smoothness condition in (5) requires connected nodes to have similar values (according to (4)), and forces the graph spectrum of the graph signal to be concentrated in the small eigenvalues region (according to (4)).

A graph filter applied on a graph signal is a linear operator that satisfies [51]

$$\mathbf{a}_{out} = \mathbf{V} \text{diag}(\psi(\lambda_1), \dots, \psi(\lambda_N)) \mathbf{V}^T \mathbf{a}_{in}, \quad (6)$$

where \mathbf{a}_{out} and \mathbf{a}_{in} are the output and input graph signals, $\text{diag}(\mathbf{a})$ is a diagonal matrix in which the (n, n) th entry is a_n , and $\psi(\lambda_n)$ is the graph filter frequency response at the graph frequency λ_n , $n = 1, \dots, N$. Low-pass graph filters of order K are defined as follows [52].

Definition 1. The graph filter in (6) is a low-pass graph filter of order K with a cutoff frequency at λ_K if $\eta_K < 1$, where

$$\eta_k \triangleq \frac{\max\{|\psi(\lambda_{k+1})|, \dots, |\psi(\lambda_N)|\}}{\min\{|\psi(\lambda_1)|, \dots, |\psi(\lambda_k)|\}}, \quad k = 1, \dots, N - 1. \quad (7)$$

This definition implies that if $\eta_K < 1$, then most of the energy of the graph filter is concentrated in the first K frequency bins of the graph filter [52]. Upon passing a graph signal through $\mathbf{V} \text{diag}(\psi(\lambda_1), \dots, \psi(\lambda_N)) \mathbf{V}^T$, the high-frequency components (related to graph frequencies greater than λ_K) are attenuated relative to the low-frequency components (related to graph frequencies lower than λ_K). Accordingly, as long as the input of the filter is a “well-behaved” excitation, and does not possess strong high-pass components, the output signal is a K -low-pass graph signal [52], and thus, a *smooth* graph signal for small K as defined in (5).

2.2. DC-PF Model: State Estimation and Observability

A power system is a network of buses (generators or loads) connected by transmission lines that can be represented as an undirected weighted graph, $\mathcal{G}(\mathcal{V}, \zeta)$, where the set of nodes, \mathcal{V} , is the set of N buses, and the edge set, ζ , is the set of P transmission lines between these buses. The set of all sensor measurements is denoted by \mathcal{M} , which includes $M \triangleq 2P + N$ active power measurements at the N buses and at the bi-directional P transmission lines. According to the π -model [1], each transmission line, $(k, n) \in \zeta$, which connects buses k and n , is characterized by an admittance value, $Y_{k,n}$.

The active power and the voltages obey multivariate versions of Kirchhoff’s and Ohm’s laws that result in the nonlinear equations of the AC-PF model (see Section 5). In order to analyze the GSP properties and to simplify the presentation of the new methods, first these equations are approximated using the DC-PF model [1], in which the states are the voltage angles. Therefore, we consider first a DC-PF model with the following noisy measurements of the active power [1]:

$$\mathbf{z} = \mathbf{H}\boldsymbol{\theta} + \mathbf{e}, \tag{8}$$

where

- $\mathbf{z} = [z_1, \dots, z_M]^T \in \mathbb{R}^M$ is the active power vector.
- $\boldsymbol{\theta} = [\theta_1, \dots, \theta_N]^T \in \mathbb{R}^N$ is the system state vector, where θ_n is the voltage angle at bus n . In low-observability systems, it is more convenient to delay the assignment of the reference angle (p. 165 in [2]). Thus, $\boldsymbol{\theta}$ includes the angle of the reference bus.
- $\mathbf{e} \in \mathbb{R}^M$ is zero-mean Gaussian noise with covariance \mathbf{R} .
- $\mathbf{H} \in \mathbb{R}^{M \times N}$ is the measurements matrix, which is determined by the topology of the network, the susceptance of the transmission lines, and the meter locations [6]. In particular, the N rows of \mathbf{H} associated with the meters on the buses that measure the total power flow of the transmission lines connected to these buses together create the nodal admittance matrix \mathbf{B} (e.g., see p. 97 in [2]) with the following (k, l) -th element:

$$B_{k,l} = \begin{cases} \sum_{n \in \mathcal{N}_k} -b_{k,n}, & k = l \\ b_{k,l}, & (k,l) \in \zeta \\ 0, & \text{otherwise} \end{cases}, \quad \forall k, l = 1, \dots, N, \tag{9}$$

where \mathcal{N}_k is the set of buses connected to bus k and $b_{k,n} < 0$ is the susceptance of $(k, n) \in \zeta$, i.e., $b_{k,n}$ equals the imaginary part of $Y_{k,n}$.

The goal of DC-PF PSSE is to recover the state vector, $\boldsymbol{\theta}$, from the measurement vector, \mathbf{z} , for various monitoring purposes [1,31]. Since $\boldsymbol{\theta}$ also includes the reference bus, without loss of generality, the angle of bus 1 (the reference bus) is set to be $\theta_1 = 0$. The PSSE in this case is implemented using the following WLS estimator [1]:

$$\hat{\boldsymbol{\theta}}^{\text{WLS}} = \arg \min_{\boldsymbol{\theta} \in \mathbb{R}^N} (\mathbf{z} - \mathbf{H}\boldsymbol{\theta})^T \mathbf{R}^{-1} (\mathbf{z} - \mathbf{H}\boldsymbol{\theta}) \tag{10}$$

such that $\theta_1 = 0$.

The solution of (10) is

$$\begin{cases} \hat{\boldsymbol{\theta}}_{\bar{\mathcal{V}}}^{\text{WLS}} = \mathbf{K}\mathbf{z} \\ \hat{\theta}_1^{\text{WLS}} = 0 \end{cases}, \tag{11}$$

where

$$\mathbf{K} \triangleq (\mathbf{H}_{\mathcal{M}, \bar{\mathcal{V}}}^T \mathbf{R}^{-1} \mathbf{H}_{\mathcal{M}, \bar{\mathcal{V}}})^{-1} \mathbf{H}_{\mathcal{M}, \bar{\mathcal{V}}}^T \mathbf{R}^{-1} \tag{12}$$

and $\bar{\mathcal{V}} \triangleq \mathcal{V} \setminus 1$ is the set of all buses except the reference bus.

For state estimation to be feasible, one needs to have enough measurements such that the system state can be uniquely determined by the WLS estimation approach. This observability requirement, before the assignment of reference angles, can be defined by one of the following (p. 165 in [2]).

Definition 2. Assume the DC-PF model from (8). The network is observable if any matrix that is obtained from \mathbf{H} by deleting one of its columns has a full column rank of $N - 1$. Alternatively, the network is observable if the following holds: $\mathbf{H}\boldsymbol{\theta} = \mathbf{0}$ if, and only if, $\boldsymbol{\theta} = \alpha\mathbf{1}$, where α is an arbitrary scalar.

In particular, since according to Definition 2, $\mathbf{H}_{\mathcal{M},\bar{\mathcal{V}}}$ has full column rank for an observable system, observability ensures that $\mathbf{H}_{\mathcal{M},\bar{\mathcal{V}}}^T \mathbf{R}^{-1} \mathbf{H}_{\mathcal{M},\bar{\mathcal{V}}}$ is nonsingular, and the WLS estimator from (11) and (12) is well-defined for any observable network. Definition 2 implies the following corollary:

Corollary 1. The network is unobservable if the conditions in Definition 2 are not satisfied.

In practice, however, network observability is not always guaranteed. In such cases, the WLS estimator from (11) cannot be implemented. Even for observable systems, errors and outliers may have a disastrous effect on the state estimation. In the following, it is shown that incorporating graphical information using GSP tools improves the state estimation performance and enables estimation even in partially observable systems.

3. GSP Properties of the States

The power system can be represented as an undirected weighted graph, $\mathcal{G}(\mathcal{V}, \xi)$, as described at the beginning of Section 2.2. In this context, the state vector, $\boldsymbol{\theta} \in \mathbb{R}^N$, and the subvector of \mathbf{z} from (8) that contains the N active power injection measurements at the N buses, denoted as \mathbf{z}_{bus} , can be interpreted as *graph signals*. In this graph representation, the nodal admittance matrix from (9) is a Laplacian matrix:

$$\mathbf{L} = \mathbf{B}. \quad (13)$$

In this section, the graph low-pass nature and the smoothness of the state vector is established in power systems under normal operation conditions, and where the Laplacian matrix is set to be the nodal admittance matrix, as defined in (13). That is, using the smoothness defined in (4) and (5), we show that

$$\mathcal{E}_{\mathbf{L}}(\boldsymbol{\theta}) = \boldsymbol{\theta}^T \mathbf{L} \boldsymbol{\theta} \leq \varepsilon, \quad (14)$$

where ε is small relative to the other parameters in the system. These results are consistent with the low-pass graph nature of the complex voltages described in [40]. It can be seen from (14) that the smoothness property depends on the system states and the network topology. The lack of measurements does not change the physical behavior; therefore, the smoothness property is not affected by the system observability.

3.1. Theoretical Validation—Output of a Low-Pass Graph Filter

First, we show analytically that the state vector is a low-pass graph signal. By substituting (13) in the model in (8), after taking only the power injection measurements, one obtains

$$\mathbf{z}_{\text{bus}} = \mathbf{L}\boldsymbol{\theta} + \mathbf{e}_{\text{bus}}, \quad (15)$$

where \mathbf{e}_{bus} contains the elements of the noise vector, \mathbf{e} , that are related to the N power measurements at the N buses. Equation (15) implies that since \mathbf{L} is a Laplacian matrix, it satisfies $\mathbf{L}\mathbf{1} = \mathbf{0}$ [53], where $\mathbf{1}$ is a vector of ones with appropriate dimensions. Therefore, the states can be recovered from (15) up to a constant shift, which can be written as

$$\theta = \mathbf{L}^\dagger(\mathbf{z}_{\text{bus}} - \mathbf{e}_{\text{bus}}) + c_1 \mathbf{1} = \mathbf{V} \Lambda^\dagger \mathbf{V}^T (\mathbf{z}_{\text{bus}} - \mathbf{e}_{\text{bus}}) + c_1 \mathbf{1}, \tag{16}$$

where c_1 is an arbitrary constant that represents the constant invariant property of the state vector [1], and the last equation is obtained by substituting (2). Without loss of generality, the value of c_1 is set to be

$$c_1 = c_2 \mathbf{1}^T (\mathbf{z}_{\text{bus}} - \mathbf{e}_{\text{bus}}), \tag{17}$$

where c_2 is an arbitrary constant. Using (17) and the definition of the pseudo-inverse operator, the model in (16) can be written as the following linear graph filter input–output model:

$$\theta = \mathbf{V} \text{diag}(\psi(\lambda_1), \dots, \psi(\lambda_N)) \mathbf{V}^T (\mathbf{z}_{\text{bus}} - \mathbf{e}_{\text{bus}}), \tag{18}$$

where

$$\psi(\lambda_n) = \begin{cases} \frac{1}{\lambda_n}, & n = 2, \dots, N \\ \frac{1}{Nc_2}, & n = 1 \end{cases}. \tag{19}$$

That is, θ is an output of a graph filter with the graph frequency response in (19). This representation holds under the assumption that the network is connected. Therefore, λ_1 is the only zero eigenvalue of \mathbf{L} with the eigenvector $\frac{1}{\sqrt{N}} \mathbf{1}$ [53].

Since the eigenvalues of \mathbf{L} are ordered, $0 = \lambda_1 < \lambda_2 \leq \lambda_3 \leq \dots \leq \lambda_N$, it can be seen that the graph frequency response in (19) decreases as n increases, as long as $c_2 > \frac{1}{N\lambda_2}$. By substituting (19) in (7), one obtains

$$\eta_k = \begin{cases} \frac{\lambda_k}{\lambda_{k+1}}, & 2 \leq k \leq N - 1 \\ \frac{1}{Nc_2\lambda_2}, & k = 1 \end{cases}, \tag{20}$$

where $\eta_k < 1$, $k = 2, \dots, N - 1$. Choosing $c_2 \gg \frac{1}{N\lambda_2}$ implies that $\eta_1 \ll 1$. Hence, according to Definition 1, $\mathbf{V} \text{diag}(\psi(\lambda_1), \dots, \psi(\lambda_N)) \mathbf{V}^T$ in (18) is a graph low-pass filter of any order $K \geq 1$. Since \mathbf{z}_{bus} in (16) includes the generated powers and loads, it can be assumed to be random [54,55], and thus, the input signal, $\mathbf{z}_{\text{bus}} - \mathbf{e}_{\text{bus}}$, does not possess strong high-pass components. Hence, as explained after Definition 1, the state vector θ is a first-order low-pass graph signal, and a smooth graph signal, as defined in (14).

3.2. Experimental Validation in IEEE Systems

In the following, the smoothness of the state signal is demonstrated in the graph frequency domain for the IEEE test case systems [56]. We also demonstrate the smoothness of the voltage magnitude vector, $\mathbf{v} \triangleq [|v_1|, \dots, |v_N|]^T$, which can be interpreted as *graph signals*, which will be used in Section 5. Figure 1 compares the normalized state vector, $\frac{\theta}{\|\theta\|}$, and its GFT (calculated using (3)), $\frac{\hat{\theta}}{\|\hat{\theta}\|}$, versus bus or spectral indices, for the IEEE 118-bus system [56]. Similarly, Figure 2 presents the normalized voltage magnitude vector, $\frac{\mathbf{v}}{\|\mathbf{v}\|}$, and its GFT, $\frac{\hat{\mathbf{v}}}{\|\hat{\mathbf{v}}\|}$, and Figure 3 presents the normalized power vector, $\frac{\mathbf{z}_{\text{bus}}}{\|\mathbf{z}_{\text{bus}}\|}$, and its GFT, $\frac{\hat{\mathbf{z}}_{\text{bus}}}{\|\hat{\mathbf{z}}_{\text{bus}}\|}$. For the sake of clarity, the vectors in Figures 1–3 have been decimated by a factor of 3.

It can be seen that most of the energy of the state signal, i.e., the phases (Figure 1) and the magnitudes (Figure 2) of the voltages, is concentrated in the low graph frequencies region. Accordingly, it can be concluded that the state vector and the voltage magnitude vector are smooth graph signals in the sense of (4). In contrast, the energy of the power injection measurement vector (Figure 3) is uniformly distributed across all graph frequencies. Thus, the power signal cannot be considered to be smooth. Similar results were obtained for other IEEE systems.

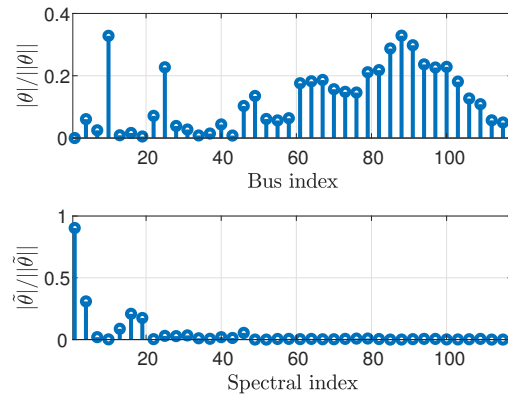


Figure 1. The state vector (**top**) and its GFT (**bottom**) for the IEEE 118-bus system.

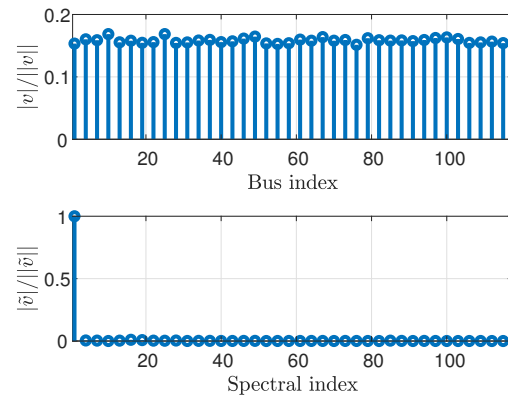


Figure 2. The voltage magnitude vector (**top**) and its GFT (**bottom**) for the IEEE 118-bus system.

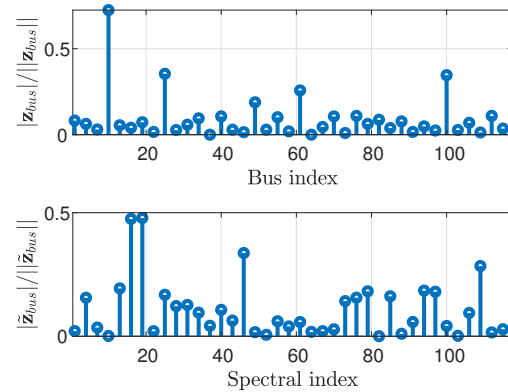


Figure 3. The active power measurement vector (**top**) and its GFT (**bottom**) for the IEEE 118-bus system.

Next, we validate experimentally that the states, θ , and the magnitudes, \mathbf{v} , are significantly smoother than the power vector, \mathbf{z}_{bus} , by comparing their normalized Dirichlet energy for typical IEEE systems, as shown in Table 1. The values of the nodal admittance matrix, $\mathbf{B} = \mathbf{L}$, the voltages, and the power data are taken from [56]. It can be seen that the phases and magnitudes are much smoother than the power injection vectors. This result is reasonable, since the phase differences between connected buses are small under normal conditions and the magnitudes are approximately constant [31], while the power may be very different, since each generator/load injects different amounts of power into the system.

Table 1. Normalized smoothness values of IEEE systems.

Measure	IEEE Test Case System				
	14-Bus	30-Bus	57-Bus	118-Bus	300-Bus
$\frac{\mathcal{E}_L(\boldsymbol{\theta})}{\ \boldsymbol{\theta}\ ^2}$	0.6617	0.3015	0.3714	1.1740	1.2371
$\frac{\mathcal{E}_L(\mathbf{v})}{\ \mathbf{v}\ ^2}$	0.0036	0.0022	0.008	0.0082	0.0199
$\frac{\mathcal{E}_L(\mathbf{z}_{\text{bus}})}{\ \mathbf{z}_{\text{bus}}\ ^2}$	16.4079	18.3307	50.8035	56.1047	138.8024

4. GSP-WLS Estimator in DC-PF Model

The recovery of smooth graph signals by incorporating regularization terms has been well studied in the GSP literature [51,57] and in the context of Laplacian regularization [58,59]. In this section, we cast the state estimation problem as a regularized graph signal recovery problem. In particular, we exploit the smoothness of the state vector, established in Section 3, to develop the smoothness-based regularized GSP-WLS estimator of the states in Section 4.1. The properties of the proposed approach are discussed in Section 4.2, where the main advantage is that it does not require system observability. In Section 4.3, an estimator of the missing power data is introduced as a by-product of this approach. Finally, in Section 4.5, a sensor allocation policy that aims to optimize the performance of the GSP-WLS estimator is designed.

4.1. GSP-WLS Estimator for the Partial Measurement Model

In the following, the case where only partial observations of the signal \mathbf{z} from (8) are available over a subset of sensors from \mathcal{M} is considered, where this subset is denoted by \mathcal{S} and $\mathcal{S} \subseteq \mathcal{M}$. A sensor at a particular location provides one row in the measurements matrix, \mathbf{H} . Therefore, based on the model in (8), the partial measurement vector can be written as

$$\mathbf{z}_{\mathcal{S}} = \mathbf{H}_{\mathcal{S},\mathcal{V}}\boldsymbol{\theta} + \mathbf{e}_{\mathcal{S}}. \quad (21)$$

Since $\mathbf{e}_{\mathcal{S}}$ contains the elements of the noise vector, \mathbf{e} , of the set of available measurements, \mathcal{S} , it is a zero-mean Gaussian noise vector with a covariance matrix $\mathbf{R}_{\mathcal{S}}$. If the columns of $\mathbf{H}_{\mathcal{S},\mathcal{V}}$ after deleting one column are linearly dependent, then, from Corollary 1, the new model in (21) with $\mathbf{H}_{\mathcal{S},\mathcal{V}}$ is not fully observable. In this *partially observable* case, the WLS estimator for the model in (21) cannot be developed via a similar strategy to that in (10) and (11), since, according to Definition 2, the state, $\boldsymbol{\theta}$, cannot be uniquely (up to a constant) determined from (21).

As a result, we need to incorporate additional properties beyond the power flow equations in (21) to obtain a valid state estimation. Here, we propose to recover $\boldsymbol{\theta}$ using the GSP-WLS estimator that incorporates the smoothness constraint from (14). Hence, the GSP-WLS estimator is defined by

$$\hat{\boldsymbol{\theta}}^{\text{GSP-WLS}} = \arg \min_{\boldsymbol{\theta} \in \mathbb{R}^N} (\mathbf{z}_{\mathcal{S}} - \mathbf{H}_{\mathcal{S},\mathcal{V}}\boldsymbol{\theta})^T \mathbf{R}_{\mathcal{S}}^{-1} (\mathbf{z}_{\mathcal{S}} - \mathbf{H}_{\mathcal{S},\mathcal{V}}\boldsymbol{\theta})$$

such that 1) $\theta_1 = 0$ and 2) $\boldsymbol{\theta}^T \mathbf{L} \boldsymbol{\theta} \leq \varepsilon$. (22)

Using $\bar{\mathcal{V}} = \mathcal{V} \setminus 1$, $\boldsymbol{\theta}_{\bar{\mathcal{V}}}$ is the state vector without the reference bus state, θ_1 , and $\mathbf{L}_{\bar{\mathcal{V}}}$ is the submatrix of \mathbf{L} obtained by removing its first row and column. The smoothness constraint in (14), after substituting $\theta_1 = 0$, can be rewritten as

$$\boldsymbol{\theta}_{\bar{\mathcal{V}}}^T \mathbf{L}_{\bar{\mathcal{V}}} \boldsymbol{\theta}_{\bar{\mathcal{V}}} \leq \varepsilon. \quad (23)$$

Using the smoothness constraint from (23) and substituting $\theta_1 = 0$ in (22), the GSP-WLS estimator is given by

$$\hat{\theta}_{\bar{v}}^{\text{GSP-WLS}} = \arg \min_{\theta_{\bar{v}} \in \mathbb{R}^{N-1}} (\mathbf{z}_{\mathcal{S}} - \mathbf{H}_{\mathcal{S},\bar{v}} \theta_{\bar{v}})^T \mathbf{R}_{\mathcal{S}}^{-1} (\mathbf{z}_{\mathcal{S}} - \mathbf{H}_{\mathcal{S},\bar{v}} \theta_{\bar{v}})$$

$$\text{such that } \theta_{\bar{v}}^T \mathbf{L}_{\bar{v}} \theta_{\bar{v}} \leq \varepsilon, \tag{24}$$

and $\hat{\theta}_1^{\text{GSP-WLS}} = 0$. Then, using the Karush–Kuhn–Tucker (KKT) conditions [60], the minimization problem in (24) can be replaced by the following regularized optimization problem (e.g., see pp. 17–19 in [61]):

$$\hat{\theta}_{\bar{v}}^{\text{GSP-WLS}} = \arg \min_{\theta_{\bar{v}} \in \mathbb{R}^{N-1}} \{(\mathbf{z}_{\mathcal{S}} - \mathbf{H}_{\mathcal{S},\bar{v}} \theta_{\bar{v}})^T \mathbf{R}_{\mathcal{S}}^{-1} (\mathbf{z}_{\mathcal{S}} - \mathbf{H}_{\mathcal{S},\bar{v}} \theta_{\bar{v}}) + \mu \theta_{\bar{v}}^T \mathbf{L}_{\bar{v}} \theta_{\bar{v}}\}, \tag{25}$$

and $\hat{\theta}_1^{\text{GSP-WLS}} = 0$. The term $\theta_{\bar{v}}^T \mathbf{L}_{\bar{v}} \theta_{\bar{v}}$ is a regularization term, which is based on the smoothness constraint from (23). The parameter $\mu \geq 0$ is a Lagrange multiplier, which is a tuning parameter that replaces ε and is discussed in Section 4.2. If the system is not fully observable based on the sensors at \mathcal{S} , then μ should be larger than zero, i.e., $\mu > 0$.

The GSP-WLS estimator from (25) is obtained by equating the derivative of (25) with respect to $\theta_{\bar{v}}$ to zero, which results in [61]

$$\begin{cases} \hat{\theta}_{\bar{v}}^{\text{GSP-WLS}} = \tilde{\mathbf{K}}(\mathcal{S}, \mu) \mathbf{z}_{\mathcal{S}} \\ \hat{\theta}_1^{\text{GSP-WLS}} = 0 \end{cases}, \tag{26}$$

where

$$\tilde{\mathbf{K}}(\mathcal{S}, \mu) \triangleq (\mathbf{H}_{\mathcal{S},\bar{v}}^T \mathbf{R}_{\mathcal{S}}^{-1} \mathbf{H}_{\mathcal{S},\bar{v}} + \mu \mathbf{L}_{\bar{v}})^{-1} \mathbf{H}_{\mathcal{S},\bar{v}}^T \mathbf{R}_{\mathcal{S}}^{-1}. \tag{27}$$

For a partially observable system, the matrix $\mathbf{H}_{\mathcal{S},\bar{v}}^T \mathbf{R}_{\mathcal{S}}^{-1} \mathbf{H}_{\mathcal{S},\bar{v}}$ is a singular matrix and the additional term in (27), and $\mu \mathbf{L}_{\bar{v}}$ with $\mu > 0$ enables the matrix inversion and improves the numerical stability of the proposed GSP-WLS estimator, since $\mathbf{L}_{\bar{v}}$ has full rank (see Lemma 1 in [62]).

4.2. Remarks

The main advantage of the proposed GSP-WLS estimator in (26) and (27) is that it does not require full observability of the system. This estimator is a function of the regularization parameter, $\mu \geq 0$. The determination of μ is discussed in Section 6. More strategies to choose μ are described in the literature (e.g., see Section 1.8 in [61]). The following are special cases of the proposed GSP-WLS estimator:

1. Full observability: If all sensors are available, then, by substituting $\mathcal{S} = \mathcal{M}$ and $\mu = 0$ in (27) one obtains that

$$\tilde{\mathbf{K}}(\mathcal{M}, 0) = \mathbf{K}, \tag{28}$$

where \mathbf{K} is defined in (12). Therefore, if $\mathcal{S} = \mathcal{M}$, then the GSP-WLS estimator from (26) with $\mu = 0$ coincides with the WLS estimator, $\hat{\theta}^{\text{WLS}}$, from (11).

2. Large μ : At the other extreme, for $\mu \rightarrow \infty$, the coefficient matrix from (27) satisfies $\lim_{\mu \rightarrow \infty} \tilde{\mathbf{K}}(\mathcal{S}, \mu) = \mathbf{0}$. Thus, in this case, the GSP-WLS estimator from (26) satisfies $\hat{\theta}^{\text{GSP-WLS}} \rightarrow \mathbf{0}$. This zero estimator can be interpreted as the a priori state estimator, which does not use the observations. Thus, taking too large a value of μ is unhelpful.
3. Relation with the pseudo-measurement WLS (pm-WLS) estimator: The pm-WLS estimator for systems that are not fully observable is based on generating pseudo-measurements of typical power injection/consumption values from historical data [1,24]. In this case, the received measurements are processed together with a priori estimated (predicted) states (without the reference bus), $\hat{\theta}_{\text{prior}} \in \mathbb{R}^{N-1}$, which are assumed to have the error covariance matrix, $\mathbf{R}_{\text{prior}} \in \mathbb{R}^{(N-1) \times (N-1)}$. The pm-WLS estimator is the maximum a posteriori state estimator [11]:

$$\begin{cases} \hat{\theta}_{\bar{v}}^{(\text{pm-WLS})} = \mathbf{K}_1 \mathbf{z}_{\mathcal{S}} + \mathbf{K}_2 \hat{\theta}_{\text{prior}} \\ \hat{\theta}_1 = 0 \end{cases}, \tag{29}$$

where

$$\mathbf{K}_1 = (\mathbf{H}_{\mathcal{S},\mathcal{V}}^T \mathbf{R}_{\mathcal{S}}^{-1} \mathbf{H}_{\mathcal{S},\mathcal{V}} + \mathbf{R}_{prior}^{-1})^{-1} \mathbf{H}_{\mathcal{S},\mathcal{V}}^T \mathbf{R}_{\mathcal{S}}^{-1} \tag{30}$$

$$\mathbf{K}_2 = (\mathbf{H}_{\mathcal{S},\mathcal{V}}^T \mathbf{R}_{\mathcal{S}}^{-1} \mathbf{H}_{\mathcal{S},\mathcal{V}} + \mathbf{R}_{prior}^{-1})^{-1} \mathbf{R}_{prior}^{-1} \tag{31}$$

It can be seen that if $\hat{\boldsymbol{\theta}}_{prior} = \mathbf{0}$ and $\mathbf{R}_{prior}^{-1} = \mu \mathbf{L}_{\mathcal{V}}$, then $\mathbf{K}_1 = \tilde{\mathbf{K}}(\mathcal{S}, \mu)$ and the pm-WLS estimator in (29) coincides with the GSP-WLS estimator in (26). Therefore, the proposed GSP-WLS estimator can be interpreted as a special case of the pm-WLS estimator, where the GSP theory provides a mathematical strategy to determine the pseudo-data information. Moreover, in general, the GSP-WLS estimator only requires setting a single scalar parameter, μ , compared with the pm-WLS estimator, which requires setting both \mathbf{R}_{prior} and $\hat{\boldsymbol{\theta}}_{prior}$.

4.3. Estimation of Missing Power Measurements

An important by-product of the GSP-WLS estimator is the following method for reconstructing the missing data of active power measurements. In the partially observable system, we have measurements obtained from the set of sensors, \mathcal{S} , which is given by $\mathbf{z}_{\mathcal{S}}$. Our goal in this subsection is to recover the other measurements that are included in the vector $\mathbf{z}_{\mathcal{M}\setminus\mathcal{S}}$. Based on the model in (8) (similar to (21)), the measurement vector of the partially observable system can be written as

$$\mathbf{z}_{\mathcal{M}\setminus\mathcal{S}} = \mathbf{H}_{\mathcal{M}\setminus\mathcal{S},\mathcal{V}} \boldsymbol{\theta} + \mathbf{e}_{\mathcal{M}\setminus\mathcal{S}}, \tag{32}$$

where $\mathbf{e}_{\mathcal{M}\setminus\mathcal{S}}$ is a zero-mean noise vector with a covariance matrix $\mathbf{R}_{\mathcal{M}\setminus\mathcal{S}}$. By substituting the GSP-WLS estimator from (26) in (32) and removing the noise term, the following WLS-type estimator of the missing power measurements is obtained:

$$\hat{\mathbf{z}}_{\mathcal{M}\setminus\mathcal{S}} = \mathbf{H}_{\mathcal{M}\setminus\mathcal{S},\mathcal{V}} \hat{\boldsymbol{\theta}}^{GSP-WLS}. \tag{33}$$

By recovering the lost power data, the EMS can also monitor the unobservable part of the system [63].

4.4. Detection of Bad Data in Unobservable Systems

State estimators can also be used for bad data detection by plugging it into any detector that is based on a state estimator. In particular, in this paper, the following bad data detection methods that are based on a general estimator $\hat{\boldsymbol{\theta}}$ can be used with the new estimator:

- Largest normalized residual test (LNR) [1]:

$$\frac{1}{\sigma^2} \|\mathbf{z}_{\mathcal{S}} - \mathbf{H}_{\mathcal{V},\mathcal{S}} \hat{\boldsymbol{\theta}}\|_{\infty} \underset{H_0}{\overset{H_1}{\geq}} \tau, \tag{34}$$

where the infinity norm of a vector \mathbf{a} is defined by $\|\mathbf{a}\|_{\infty} = \max_i |a_i|$.

- $J(\boldsymbol{\theta})$ test with $\mathbf{R} = \sigma^2 \mathbf{I}$ [1]:

$$J(\hat{\boldsymbol{\theta}}) = \frac{1}{\sigma^2} (\mathbf{z}_{\mathcal{S}} - \mathbf{H}_{\mathcal{V},\mathcal{S}} \hat{\boldsymbol{\theta}})^T (\mathbf{z}_{\mathcal{S}} - \mathbf{H}_{\mathcal{V},\mathcal{S}} \hat{\boldsymbol{\theta}}) \underset{H_0}{\overset{H_1}{\geq}} \tau. \tag{35}$$

- The GFT-based detector from [7] that was developed for the detection of false data injection (FDI) attacks. The GFT-based detection scheme calculates the GFT of an estimated grid state, $\hat{\boldsymbol{\theta}}$, and filters the graph’s high-frequency components. By comparing the maximum norm of this outcome with a threshold, it can detect the presence of FDI attacks.

In both (34) and (35), τ denotes a chosen threshold, and \mathcal{H}_0 and \mathcal{H}_1 denote the hypotheses of good/bad data, respectively.

Then, the proposed estimator from (26) can be used for bad data detection by plugging it into the detectors that are based on a state estimator; that is, by substituting $\hat{\theta} = \hat{\theta}^{\text{GSP-WLS}}$ in the LNR test from (34), the $J(\theta)$ test from (35), and the GFT-based detector from [7].

4.5. Optimization of the Sampling Policy

Sensor locations have a significant impact on the estimation performance of power systems [64]. Therefore, in this subsection, we design a sensor allocation policy for the model in (21) that aims to minimize the MSE of the GSP-WLS estimator, $\text{MSE}(\hat{\theta}) = \text{E}[(\hat{\theta} - \theta)^T(\hat{\theta} - \theta)]$. However, it can be shown that the MSE of $\hat{\theta}^{\text{GSP-WLS}}$ is a function of the unknown state vector, θ , and thus, cannot be used as an objective function for the optimization of the sensor locations. Therefore, the MSE is replaced by the CRB [65], which is a lower bound on the MSE.

In this subsection, we treat the MSE, bias, and CRB of the vector $\theta_{\bar{v}}$ (without the reference bus for the sake of simplicity). By substituting $\theta_1 = 0$ in the model in (21), one obtains that the partial measurement vector obtained from a sensor subset \mathcal{S} is a Gaussian vector with mean $\mathbf{H}_{\mathcal{S},\bar{v}}\theta_{\bar{v}}$ and covariance $\mathbf{R}_{\mathcal{S}}$:

$$\mathbf{z}_{\mathcal{S}} \sim \mathcal{N}(\mathbf{H}_{\mathcal{S},\bar{v}}\theta_{\bar{v}}, \mathbf{R}_{\mathcal{S}}). \tag{36}$$

The CRB for this Gaussian vector, which is a lower bound on the MSE, is given by (pp. 45–46 in [65])

$$\text{CRB}(\mathcal{S}) \triangleq \text{Tr} \left(\left(\mathbf{I} + \frac{\partial \mathbf{b}(\mathcal{S})}{\partial \theta_{\bar{v}}} \right) (\mathbf{H}_{\mathcal{S},\bar{v}}^T \mathbf{R}_{\mathcal{S}}^{-1} \mathbf{H}_{\mathcal{S},\bar{v}})^{\dagger} \left(\mathbf{I} + \frac{\partial \mathbf{b}(\mathcal{S})}{\partial \theta_{\bar{v}}} \right)^T \right), \tag{37}$$

where $\mathbf{b}(\mathcal{S}) \triangleq \text{E}[\hat{\theta}_{\bar{v}} - \theta_{\bar{v}}]$ is the bias of the estimator and $\frac{\partial \mathbf{b}(\mathcal{S})}{\partial \theta_{\bar{v}}}$ is its gradient. Using the model in (21) and the estimator in (26), it can be seen that the bias of the GSP-WLS estimator is

$$\mathbf{b}(\mathcal{S}) = \text{E}[\hat{\theta}_{\bar{v}}^{\text{GSP-WLS}} - \theta_{\bar{v}}] = \tilde{\mathbf{K}}(\mathcal{S}, \mu) \mathbf{H}_{\mathcal{S},\bar{v}} \theta_{\bar{v}} - \theta_{\bar{v}}, \tag{38}$$

where $\tilde{\mathbf{K}}(\mathcal{S}, \mu)$ is defined in (27). Accordingly, the gradient of (38) with respect to $\theta_{\bar{v}}$ is

$$\frac{\partial \mathbf{b}(\mathcal{S})}{\partial \theta_{\bar{v}}} = \tilde{\mathbf{K}}(\mathcal{S}, \mu) \mathbf{H}_{\mathcal{S},\bar{v}} - \mathbf{I}. \tag{39}$$

By substituting (39) in (37), we obtain that the CRB on the MSE of estimators with the GSP-WLS bias is given by

$$\text{CRB}(\mathcal{S}) = \text{Tr} \left(\tilde{\mathbf{K}}(\mathcal{S}, \mu) \mathbf{H}_{\mathcal{S},\bar{v}} (\mathbf{H}_{\mathcal{S},\bar{v}}^T \mathbf{R}_{\mathcal{S}}^{-1} \mathbf{H}_{\mathcal{S},\bar{v}})^{\dagger} \mathbf{H}_{\mathcal{S},\bar{v}}^T \tilde{\mathbf{K}}^T(\mathcal{S}, \mu) \right). \tag{40}$$

By substituting (27) in (40) and using the pseudo-inverse property, $\mathbf{A} = \mathbf{A}\mathbf{A}^{\dagger}\mathbf{A}$, one obtains

$$\text{CRB}(\mathcal{S}) = \text{Tr}(\tilde{\mathbf{K}}(\mathcal{S}, \mu) \mathbf{R}_{\mathcal{S}} \tilde{\mathbf{K}}^T(\mathcal{S}, \mu)). \tag{41}$$

The CRB in (41) is not a function of the unknown state vector, θ , and thus can be used as an optimization criterion for choosing the sensor locations. We assume a constrained quantity of sensing resources, e.g., due to a limited energy and communication budget. Therefore, the problem of the selection of sensor locations with only \tilde{q} sensors can be written as follows:

$$\mathcal{S}^{\text{opt}} = \arg \min_{\mathcal{S} \subset \mathcal{M}} \text{CRB}(\mathcal{S}) \text{ s.t. } |\mathcal{S}| = \tilde{q} = \arg \min_{\mathcal{S} \subset \mathcal{M}} (\tilde{\mathbf{K}}(\mathcal{S}, \mu) \mathbf{R}_{\mathcal{S}} \tilde{\mathbf{K}}^T(\mathcal{S}, \mu)) \text{ s.t. } |\mathcal{S}| = \tilde{q}, \tag{42}$$

where the last equality is obtained by substituting (41).

It is assumed that in the measured buses, all the relevant power measurements are given. Thus, \mathcal{S} is uniquely determined by the buses chosen for the measurements. For the sake of simplicity, in the optimization approach, we take $\mathbf{H}_{\mathcal{V},\bar{\mathcal{V}}} = \mathbf{L}_{\mathcal{V},\bar{\mathcal{V}}}$ and replace the selection of \tilde{q} sensors by the selection of q buses. Therefore, by substituting $\mathbf{H}_{\mathcal{V},\bar{\mathcal{V}}} = \mathbf{L}_{\mathcal{V},\bar{\mathcal{V}}}$, the problem in (42) is replaced by the problem of selecting the optimal buses in the CRB sense. However, finding the set of q locations among all the N buses with the smallest CRB is a combinatorial optimization with a computational complexity of $\binom{N}{q}$, which is practically infeasible. Therefore, we propose a greedy algorithm, Algorithm 1, for the practical implementation of the sampling scheme. The idea behind this algorithm is to iteratively add to the sampling set those buses that lead to the minimal CRB. In addition, Figure 4 illustrates the data flow diagram of this greedy algorithm for selecting the measured buses.

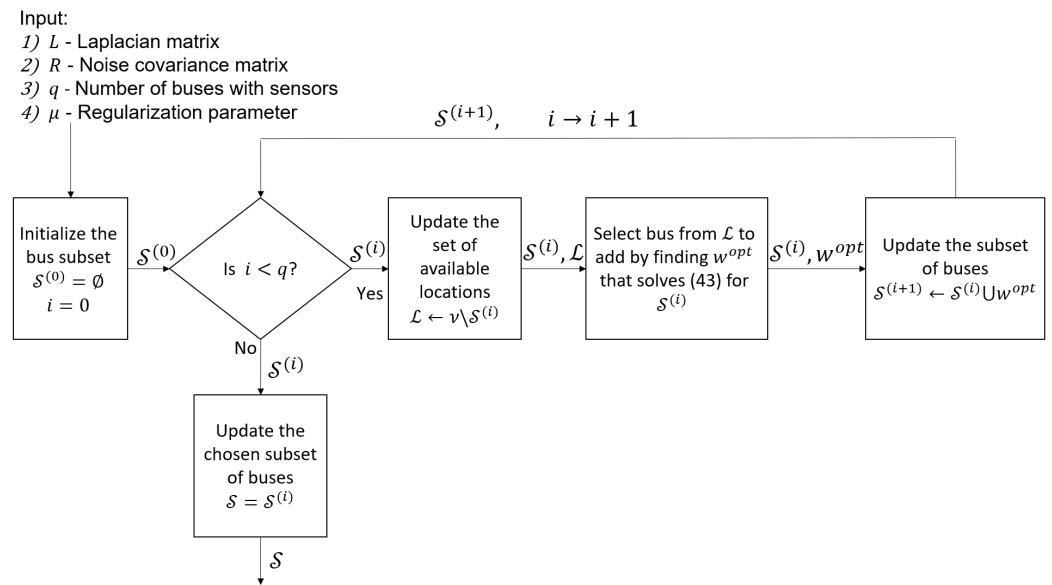


Figure 4. Flow of the proposed GSP greedy selection of the measured buses (Algorithm 1).

Algorithm 1 Greedy selection of the measured buses

Input:

- (1) Laplacian matrix, \mathbf{L} , and noise covariance matrix, \mathbf{R}
- (2) Number of buses with sensors, q
- (3) Regularization parameter, μ

Output: Subset of q buses, \mathcal{S}

- 1: Initialize the bus subset $\mathcal{S}^{(0)} = \emptyset$ and the iteration, $i = 0$
- 2: **while** $i < q$ **do**
- 3: Update the set of available locations, $\mathcal{L} = \mathcal{V} \setminus \mathcal{S}^{(i)}$
- 4: Find the optimal bus to add:

$$w^{opt} = \arg \min_{w \in \mathcal{L}} \text{Tr}(\{\tilde{\mathbf{K}}(\mathcal{S}^{(i)} \cup w), \mu\} \mathbf{R}_{(\mathcal{S}^{(i)} \cup w)} \tilde{\mathbf{K}}^T(\{\mathcal{S}^{(i)} \cup w\}, \mu)), \quad (43)$$

where $\tilde{\mathbf{K}}$ is defined in (27) with $\mathbf{H}_{\mathcal{V},\bar{\mathcal{V}}} = \mathbf{L}_{\mathcal{V},\bar{\mathcal{V}}}$

- 5: Update the subset of buses, $\mathcal{S}^{(i+1)} \leftarrow \mathcal{S}^{(i)} \cup w^{opt}$, and the iteration, $i \leftarrow i + 1$
 - 6: **end while**
 - 7: Update the chosen subset of buses: $\mathcal{S} = \mathcal{S}^{(i)}$
-

5. Extension to the AC-PF Model

Since the problem of low observability mainly occurs in distribution systems, which requires AC state estimation, in this section, the GSP-WLS estimator is extended to the AC-PF model, where voltage magnitudes are estimated as well. In particular, the conventional

PSSE in the AC-PF model is described in Section 5.1. Then, in Section 5.2, the proposed iterative regularized Gauss–Newton method that exploits the smoothness property of the voltage phases and magnitudes in each iteration is presented. In Section 5.3, the properties of the proposed GSP Gauss–Newton algorithm are discussed.

5.1. Model, State Estimation, and Observability

In the following, we replace the DC-PF model from (8) by the following nonlinear AC-PF model equations:

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \mathbf{e}, \quad (44)$$

where

- $\mathbf{z} \in \mathbb{R}^M$ is the measurement vector that includes the active and reactive branch power flows and power injections.
- $\mathbf{h}(\mathbf{x})$ is the measurement function, which is determined by the sensor types and their locations in the network.
- $\mathbf{x} = [\theta_2, \dots, \theta_N, |v_2|, \dots, |v_N|]^T \in \mathbb{R}^{2N-2}$, is the state vector here, where bus 1 is the reference bus, and thus, $\theta_1 = 0$ and v_1 is known (e.g., see Chapter 4 in [1]).

The specific forms and parameters of (44) with different levels of modeling details can be found, e.g., in Chapter 2 of [1].

Similar to the WLS estimator for the DC-PF model in (10), the AC-PF state estimator is usually based on minimizing the following WLS objective function:

$$J(\mathbf{z}, \mathbf{h}(\mathbf{x}), \mathbf{R}) = (\mathbf{z} - \mathbf{h}(\mathbf{x}))^T \mathbf{R}^{-1} (\mathbf{z} - \mathbf{h}(\mathbf{x})), \quad (45)$$

with respect to \mathbf{x} [1]. The first-order optimality condition for the unconstrained minimization problem in (45) is given by

$$\mathbf{g}(\mathbf{x}) \triangleq \frac{\partial J(\mathbf{z}, \mathbf{h}(\mathbf{x}), \mathbf{R})}{\partial \mathbf{x}} = -\mathbf{H}^T(\mathbf{x}) \mathbf{R}^{-1} (\mathbf{z} - \mathbf{h}(\mathbf{x})) = \mathbf{0}, \quad (46)$$

where $\mathbf{H}(\mathbf{x}) \triangleq \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$ is the Jacobian matrix of $\mathbf{h}(\mathbf{x})$ at \mathbf{x} . Solving the nonlinear equation in (46) using the Gauss–Newton method [1,2] results in the following iterative system:

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} + \mathbf{G}^{-1}(\mathbf{x}^{(i)}) \mathbf{H}^T(\mathbf{x}^{(i)}) \mathbf{R}^{-1} (\mathbf{z} - \mathbf{h}(\mathbf{x}^{(i)})), \quad (47)$$

where $\mathbf{x}^{(i)}$ is the state estimator at the i th iteration and

$$\mathbf{G}(\mathbf{x}) = \mathbf{H}^T(\mathbf{x}) \mathbf{R}^{-1} \mathbf{H}(\mathbf{x}) \quad (48)$$

is the gain matrix. Iterating until convergence, i.e., until $\|\mathbf{x}^{(i+1)} - \mathbf{x}^{(i)}\| \leq \delta$, one will obtain the solution of PSSE.

The observability requirement for the AC-PF model can be defined as follows (see Chapter 4.6 in [1], Donti et al. [5]).

Definition 3. Assume the AC-PF model from (44). The network is observable if $\mathbf{G}(\mathbf{x})$ is a nonsingular matrix for any \mathbf{x} in the solution space.

By observing (48), it can be seen that if $\mathbf{H}(\mathbf{x})$ has a full column rank of $2N - 2$, then the network is observable in the AC-PF sense. This observability condition should be satisfied in each iteration of the Gauss–Newton iterative algorithm.

5.2. GSP-Based Gauss–Newton Algorithm

Similar to Section 4, here we consider the case where only partial observations of \mathbf{z} from (44) are available over a subset of sensors $\mathcal{S} \subseteq \mathcal{M}$. That is, based on the model in (44), the partial measurement AC-PF model can be written as

$$\mathbf{z}_S = \mathbf{h}_S(\mathbf{x}) + \mathbf{e}_S, \tag{49}$$

where \mathbf{e}_S is a zero-mean Gaussian noise vector with a covariance matrix \mathbf{R}_S , as in (21). The Jacobian matrix of the model in (49) is $\mathbf{H}_{S,\bar{\mathcal{V}}}(\mathbf{x}) = \frac{\partial \mathbf{h}_S(\mathbf{x})}{\partial \mathbf{x}}$, where $\bar{\mathcal{V}}$ indicates the set of all the columns in $\mathbf{H}(\mathbf{x})$. If the columns of $\mathbf{H}_{S,\bar{\mathcal{V}}}(\mathbf{x})$ are linearly dependent, then $\mathbf{G}(\mathbf{x})$ is a singular matrix, and from Definition 3, the new system in (49) is not fully observable. In this case, the Gauss–Newton iterative procedure for the minimization of (45) cannot be implemented, since the update of the solution cannot be uniquely determined from (47).

In order to tackle this problem, we incorporate the smoothness constraints from (23) and from (5) with $\mathbf{a} = \mathbf{v}$. Hence, the GSP-WLS estimator for the AC-PF model is defined by

$$\begin{aligned} \hat{\mathbf{x}}^{\text{GSP-WLS}} &= \arg \min_{\mathbf{x}=[\boldsymbol{\theta}_{\bar{\mathcal{V}}}^T, \mathbf{v}_{\bar{\mathcal{V}}}^T]^T \in \mathbb{R}^{2N-2}} J(\mathbf{z}_S, \mathbf{h}_S(\mathbf{x}), \mathbf{R}_S) \\ &\text{such that 1) } \boldsymbol{\theta}_{\bar{\mathcal{V}}}^T \mathbf{L}_{\bar{\mathcal{V}}} \boldsymbol{\theta}_{\bar{\mathcal{V}}} \leq \varepsilon_{\theta} \\ &\text{and 2) } (\mathbf{v}_{\bar{\mathcal{V}}} - v_1 \mathbf{1})^T \mathbf{L}_{\bar{\mathcal{V}}} (\mathbf{v}_{\bar{\mathcal{V}}} - v_1 \mathbf{1}) \leq \varepsilon_{\mathbf{v}}, \end{aligned} \tag{50}$$

where the function J is defined in (45), and $\varepsilon_{\theta}, \varepsilon_{\mathbf{v}}$ are the tuning parameters of the smoothness of $\boldsymbol{\theta}$ and \mathbf{v} .

Using the KKT conditions [60], the minimization in (50) can be replaced by the following regularized optimization:

$$\hat{\mathbf{x}}^{\text{GSP-WLS}} = \arg \min_{\mathbf{x} \in \mathbb{R}^{2N-2}} J_{\text{reg}}(\mathbf{z}_S, \mathbf{h}_S(\mathbf{x}), \mathbf{R}_S, \bar{\mathbf{L}}(\mu_{\theta}, \mu_{\mathbf{v}})), \tag{51}$$

where

$$J_{\text{reg}}(\mathbf{z}, \mathbf{h}(\mathbf{x}), \mathbf{R}, \bar{\mathbf{L}}(\mu_{\theta}, \mu_{\mathbf{v}})) \triangleq J(\mathbf{z}, \mathbf{h}(\mathbf{x}), \mathbf{R}) + (\mathbf{x} - \mathbf{x}_0)^T \bar{\mathbf{L}}(\mu_{\theta}, \mu_{\mathbf{v}}) (\mathbf{x} - \mathbf{x}_0), \tag{52}$$

$$\bar{\mathbf{L}}(\mu_{\theta}, \mu_{\mathbf{v}}) \triangleq \begin{bmatrix} \mu_{\theta} \mathbf{L}_{\bar{\mathcal{V}}} & \mathbf{0} \\ \mathbf{0} & \mu_{\mathbf{v}} \mathbf{L}_{\bar{\mathcal{V}}} \end{bmatrix}, \tag{53}$$

and $\mathbf{x}_0 \triangleq [\mathbf{0}^T, v_1 \mathbf{1}^T]^T$. The right term in (52) is a regularization term, which is based on the smoothness property of the phases and magnitudes of the voltages, established in Section 3. The parameters $\mu_{\theta}, \mu_{\mathbf{v}} \geq 0$ are Lagrange multipliers that replace $\varepsilon_{\theta}, \varepsilon_{\mathbf{v}}$ as regularization parameters, and their tuning is discussed in Sections 5.3 and 6.3.

The minimum of the quadratic objective function, $J_{\text{reg}}(\mathbf{z}_S, \mathbf{h}_S(\mathbf{x}), \mathbf{R}_S, \bar{\mathbf{L}}(\mu_{\theta}, \mu_{\mathbf{v}}))$, with respect to \mathbf{x} can be determined using the first order optimality conditions as follows:

$$\mathbf{g}_{\text{reg}}(\mathbf{x}) \triangleq \frac{\partial J_{\text{reg}}(\mathbf{z}_S, \mathbf{h}_S(\mathbf{x}), \mathbf{R}_S, \bar{\mathbf{L}}(\mu_{\theta}, \mu_{\mathbf{v}}))}{\partial \mathbf{x}} = -\mathbf{H}_{S,\bar{\mathcal{V}}}^T(\mathbf{x}) \mathbf{R}_S^{-1} (\mathbf{z}_S - \mathbf{h}_S(\mathbf{x})) + \bar{\mathbf{L}}(\mu_{\theta}, \mu_{\mathbf{v}}) (\mathbf{x} - \mathbf{x}_0) = \mathbf{0}. \tag{54}$$

Then, similar to (47), the nonlinear equation, $\mathbf{g}_{\text{reg}}(\mathbf{x}) = \mathbf{0}$, is solved using the following Gauss–Newton method iteration:

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} + \mathbf{G}_{\text{reg}}^{-1}(\mathbf{x}^{(i)}) (\mathbf{H}_{S,\bar{\mathcal{V}}}^T(\mathbf{x}^{(i)}) \mathbf{R}_S^{-1} (\mathbf{z}_S - \mathbf{h}_S(\mathbf{x}^{(i)})) - \bar{\mathbf{L}}(\mu_{\theta}, \mu_{\mathbf{v}}) (\mathbf{x}^{(i)} - \mathbf{x}_0)), \tag{55}$$

where

$$\mathbf{G}_{\text{reg}}(\mathbf{x}) \triangleq \frac{\partial \mathbf{g}_{\text{reg}}(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{H}_{S,\bar{\mathcal{V}}}^T(\mathbf{x}) \mathbf{R}_S^{-1} \mathbf{H}_{S,\bar{\mathcal{V}}}(\mathbf{x}) + \bar{\mathbf{L}}(\mu_{\theta}, \mu_{\mathbf{v}}) \tag{56}$$

is the new gain matrix. Solving this equation and iterating until the required accuracy, δ , is reached, i.e., $\|\mathbf{x}^{(i+1)} - \mathbf{x}^{(i)}\| \leq \delta$, one will obtain the proposed GSP-WLS estimator for the AC-PF model. It can be seen that for a partially observable system, $\mathbf{H}_{S,\bar{\mathcal{V}}}^T(\mathbf{x}) \mathbf{R}_S^{-1} \mathbf{H}_{S,\bar{\mathcal{V}}}(\mathbf{x})$ is a singular matrix, and the additional terms in (56), $\bar{\mathbf{L}}(\mu_{\theta}, \mu_{\mathbf{v}})$ from (53) with $\mu_{\theta} > 0$, and/or $\mu_{\mathbf{v}} > 0$, can enable the matrix inversion of $\mathbf{G}_{\text{reg}}(\mathbf{x})$ and improve the numerical stability

of the GSP-WLS estimator for the AC-PF model. The iterative solution is summarized in Algorithm 2. For the initialization, we suggest that one use the “flat start”, where all bus voltages are 1 per unit and have the same phase [1]. Figure 5 presents the flow of the iterative regularized Gauss–Newton algorithm through a data flow diagram.

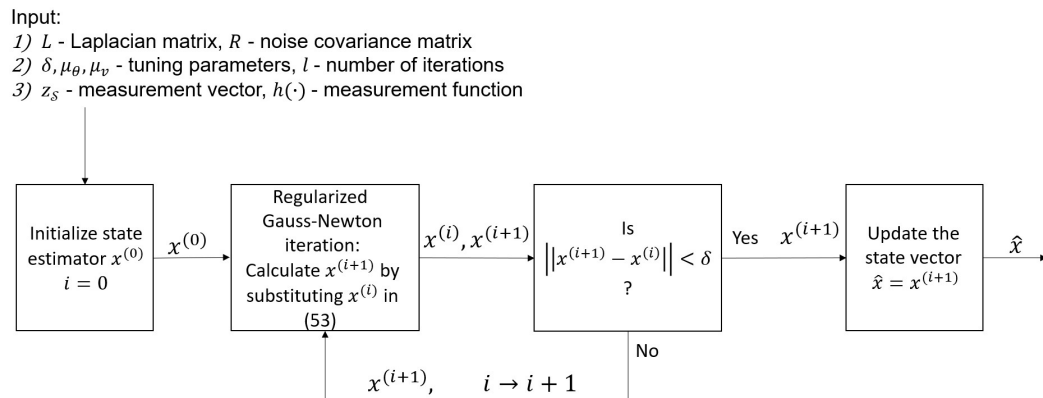


Figure 5. Flow of the proposed regularized Gauss–Newton scheme.

Algorithm 2 Regularized Gauss–Newton (GSP-WLS)

Input:

- (1) Laplacian matrix, L , and noise covariance matrix, R_S
- (2) Tuning parameters: $\delta, \mu_\theta, \mu_v$ and number of iterations, l
- (3) Measurement vector, z_S , and the function, $h_S(\cdot)$

Output: State estimator, \hat{x}

- 1: Initialize the state vector $x^{(0)}$
 - 2: **for** $i = 0, \dots, l$ **do**
 - 3: Calculate the right hand side of (55) for $x^{(i)}$
 - 4: Solve (55) for $x^{(i+1)}$
 - 5: **if** $\|x^{(i+1)} - x^{(i)}\| < \delta$: **break**
 - 6: **end for**
 - 7: Update the state vector: $\hat{x} = x^{(i+1)}$
-

5.3. Remarks

In the following, we present special cases of the proposed GSP-WLS estimator for the AC-PF model implemented by the regularized Gauss–Newton method.

(1) Full observability: If the information from all sensors is available, i.e., if $h_S(x) = h(x)$, then, by substituting $\mu_\theta = \mu_v = 0$ in (55) one obtains that $G_{reg}(x) = G(x)$, where $G(x)$ and $G_{reg}(x)$ are defined in (48) and (56), respectively. Accordingly, in this case, the Gauss–Newton iteration of the GSP-WLS estimator in (55) coincides with the Gauss–Newton iteration in (47).

(2) Relation with the pm-WLS estimator: Similar to the case of the DC-PF model, the pm-WLS estimator for partially observable systems is calculated based on the measurements and the a priori estimated (predicted) states, $\hat{x}_{prior} \in \mathbb{R}^{2N-2}$, with the forecasting error covariance matrix, R_{prior} . As a result, the following pm-WLS estimator is obtained [11]:

$$\hat{x}^{pm-WLS} = \arg \min_{x \in \mathbb{R}^{2N-2}} \left\{ J(z_S, h_S(x), R_S) + (x - \hat{x}_{prior})^T R_{prior}^{-1} (x - \hat{x}_{prior}) \right\}. \quad (57)$$

The pm-WLS estimator for the AC-PF model can be calculated with the Gauss–Newton method [11]. It can be seen that if we substitute $\hat{x}_{prior} = x_0$ and $R_{prior}^{-1} = \bar{L}(\mu_\theta, \mu_v)$, then the pm-WLS estimator from (57) coincides with the GSP-WLS estimator from (51) and (52). Hence, similar to the DC-PF model, the proposed GSP-WLS estimator can be interpreted as a special case of the pm-WLS estimator, where the GSP theory provides the mathematical justification for the determination of the pseudo-data information.

6. Results

In this section, the performance of the proposed methods is compared with that of existing methods. In Section 6.2, the performance of the GSP-WLS estimator from Section 4 is evaluated. In Section 6.3, the performance of the regularized Gauss–Newton method from Section 5 is investigated. The influence of the sampling policy from Section 4.5 is examined for both cases. In Section 6.4, the use of the new estimator for bad data detection is demonstrated.

6.1. Simulations Platform and Parameters

All the simulations were performed with Matlab R2020b. The measurements were generated according to the AC-PF model from (44) with $\mathbf{h}(\cdot)$

$$\begin{aligned} h_{2n-1}(\mathbf{x}) &= \sum_{k=1}^N |v_n||v_k|(G_{n,k} \cos(\theta_n - \theta_k) + B_{n,k} \sin(\theta_n - \theta_k)), \quad n = 1, \dots, N, \\ h_{2n}(\mathbf{x}) &= \sum_{k=1}^N |v_n||v_k|(G_{n,k} \sin(\theta_n - \theta_k) - B_{n,k} \cos(\theta_n - \theta_k)), \quad n = 1, \dots, N, \end{aligned} \quad (58)$$

for the real and reactive power injection measurement at bus n and

$$\begin{aligned} h_{2N+2f(n,k)-1}(\mathbf{x}) &= |v_n|^2 G_{n,n} - |v_k||v_n|(G_{n,k} \cos(\theta_n - \theta_k) + B_{n,k} \sin(\theta_n - \theta_k)), \quad \forall n \in \mathcal{N}_k, \quad k = 1, \dots, N, \\ h_{2N+2f(n,k)}(\mathbf{x}) &= -|v_n|^2 B_{n,n} - |v_n||v_k|(G_{n,k} \sin(\theta_n - \theta_k) - B_{n,k} \cos(\theta_n - \theta_k)), \quad \forall n \in \mathcal{N}_k, \quad k = 1, \dots, N, \end{aligned} \quad (59)$$

for the real and reactive power flow from bus n to bus k measurements ($f(n, k)$ is a one-to-one mapping of all $(n, k) \in \zeta$ to $[1, 2P]$), where $\mathbf{x} = [\theta_2, \dots, \theta_N, |v_2|, \dots, |v_N|]^T \in \mathbb{R}^{2N-2}$ is the state vector here, bus 1 is the reference bus, and thus $\theta_1 = 0$ and v_1 is known. The values of the conductance and the susceptance matrices, \mathbf{B} and \mathbf{G} , as well as the values of the voltage angles and magnitudes, were taken from the IEEE 118-bus test case recorded in [56]. This IEEE 118-bus test case represents a simple approximation of the American electric power system (in the U.S. Midwest) as of December 1962. This IEEE 118-bus system, which has $N = 118$ buses and, at most, $M = 952$ power measurements, contains 19 generators, 35 synchronous condensers, 177 lines, 9 transformers, and 91 loads. Bus number 1 is set to be the slack bus. Gaussian white noise was added to generate several measurements from the recorded grid state. In particular, in the simulations described in the following subsections, the noise covariance matrix is set to $\mathbf{R} = \sigma^2 \mathbf{I}$, where, unless otherwise stated, $\sigma^2 = 0.01$. The performance is evaluated using 1000 Monte Carlo simulations.

We compare the estimation performance of the different estimators implemented for the following bus selection policies:

- (i) Random bus selection policy (rand.)—the measured buses are randomly chosen independently from $\{1, \dots, N\}$, where for more than 72 buses only observable systems are taken.
- (ii) Experimentally designed sampling (E-design) [38]—the buses are chosen to maximize the smallest singular value of the matrix $\mathbf{V}_{S, \{1, \dots, R\}}$, where R is set to 48. The basic assumption behind this method (which was suggested in [66] for power systems) is that the measured graph signal (here, the power signal) is an R -bandlimited signal in the graph frequency domain. That is, the GFT of \mathbf{z}_{bus} satisfies $(\tilde{z}_{\text{bus}})_n = 0$, $n = R + 1, \dots, N$, where R is the cutoff frequency. As can be seen in Figure 3, in practice, the R -bandlimitness assumption does not hold for the power signal.
- (iii) Minimum CRB (Algorithm 1)—the proposed bus selection policy from Algorithm 1.

It should be noted that the CRB from (37) is not presented in the following simulations, since it is not the main goal of this study, and in order to increase the interpretability of the figures.

Figure 6 presents the estimated probability for the system to be observable, according to Definition 2, versus the number of measured buses. The estimated probability of observability

is calculated as the percentage of scenarios with observable systems in 100,000 Monte Carlo simulations for randomly selected buses in the system. It can be seen that this probability of observability increases as the number of measured buses increases, and that for fewer than 72 measured buses, the IEEE 118-bus system will not be observable with probability 1.

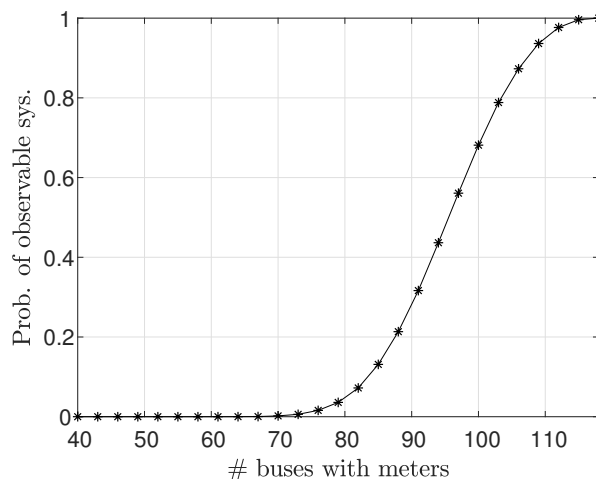


Figure 6. The estimated probability for the IEEE 118-bus system to be observable versus the number of measured buses.

6.2. State Estimation and Sampling under the DC-PF Model

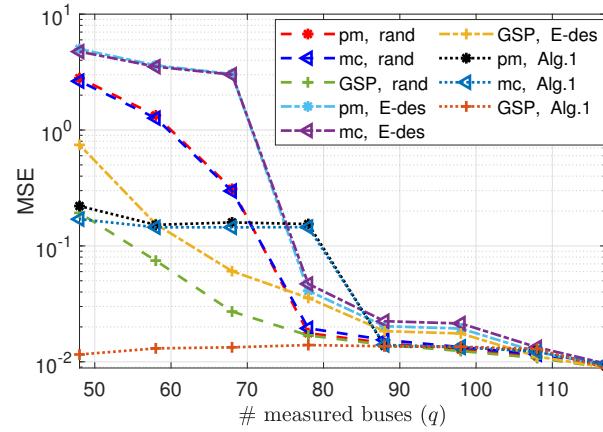
In this subsection, we evaluate the performance of the following estimators:

1. The pm-WLS estimator from [11], generated with $\mathbf{R}_{prior}^{-1} = 0.5\mathbf{I}$, and where $\hat{\theta}_{prior}$ is randomly chosen from a zero-mean Gaussian distribution with covariance $0.015\mathbf{I}$.
2. The matrix compilation (mc) method [5], implemented in this subsection by substituting the DC model from (8) in the constraints of the method in Equations (8) and (12) in [5].
3. The proposed GSP-WLS estimator (26) and (27) with $\mu = 0.1$.

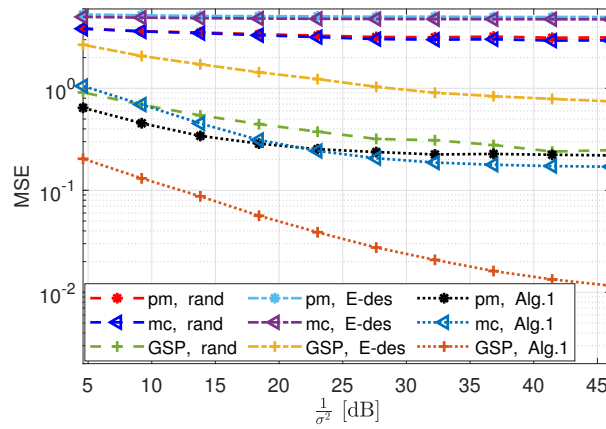
In Figure 7a,b the MSE of the GSP-WLS, pm-WLS, and mc estimators is presented versus the number of measured buses, q , and versus $\frac{1}{\sigma^2}$, respectively, with the sampling policies (i)–(iii). Figure 7b is obtained for $q = 48$ measured buses, for which the system is not observable with probability 1. It can be seen that the MSE decreases as q increases and as σ^2 decreases. In both figures, the GSP-WLS estimator outperforms the pm-WLS and mc estimators for any tested sampling policy. In this case, the performance of the pm-WLS and the mc estimators is similar, since it can be shown that for the mc estimator, the rank regularization on θ (see Eq. (13a) in [5]) turns out to be an l_2 regularization on $\theta - \hat{\theta}_{prior}$, as in (29), where $\hat{\theta}_{prior}$ approaches zero. In addition, it can be seen that the proposed sampling policy from Algorithm 1 results in a significantly lower MSE than that obtained for the random and the E-design sampling policies for all estimators. In Figure 7a, it can be seen that for each sampling policy, the MSE of the GSP-WLS, pm-WLS, and mc estimators coincides with that of the other methods where the system becomes observable (i.e., where $q > 72$ for the random sampling, $q > 76$ for the E-design sampling, and $q > 79$ for Algorithm 1). In particular, where $q = 118$ (i.e., full observability with probability 1), the MSEs of all estimators are identical.

Figure 8 shows the MSE of the power estimator \hat{z} from (33) with the three estimators and the three sampling policies. It can be seen that the MSE decreases as the number of measured buses increases, as expected, since there are fewer parameters to estimate with the increase in the number of samples and the state estimation is more accurate, as presented in Figure 7a. It can be seen that the relationships between the sampling policies and the estimators are similar to Figure 7a, where the GSP-WLS estimator with the bus selection policy of Algorithm 1 achieves the lowest MSE. Moreover, for each sampling policy, the

MSE of the power estimation based on the GSP-WLS, pm-WLS, and mc estimators coincides with that of the other methods where the system becomes observable (i.e., where $q > 72$ for the random sampling, $q > 76$ for the E-design sampling, and $q > 79$ for Algorithm 1).



(a)



(b)

Figure 7. State estimation under the DC–PF model: the MSE of the GSP–WLS from (26) and (27), pm–WLS [11], and mc [5] estimators for random, E–design [38], and Algorithm 1 bus selection policies versus (a) the number of buses, q , with $\sigma^2 = 0.01$, and (b) $\frac{1}{\sigma^2}$ with $q = 48$.

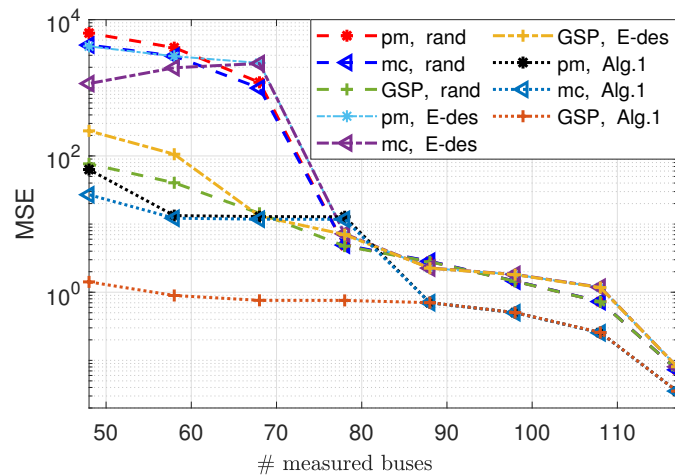


Figure 8. Power estimation based on the DC–PF model: the MSE of the GSP–WLS (26) and (27), pm–WLS [11], and mc [5] estimators for random, E–design [38], and Algorithm 1 bus selection policies versus the number of buses, q , where $\sigma^2 = 0.01$.

In Table 2 and in Figure 9, the influence of the tuning parameter, μ , is examined; we show that the proposed GSP-WLS estimator is robust to the choice of this tuning parameter by demonstrating that the estimator achieves the same performance for a range of values of μ . In Table 2, the average and the sample standard deviation of the MSE of the GSP-WLS estimator for different values of $\mu \in [0.01, 10]$ are presented, for a system with $q = 48$ and $\sigma^2 = 0.01$, for the three sampling policies. It can be seen that, for the tested scenario, the MSE is approximately constant for any μ in the range $[0.01, 10]$. Of course, further increasing μ will eventually increase the MSE, since the weight of the measurements in the estimation will be negligible. Similar results were obtained for other values of q and σ . Therefore, choosing any μ in this wide range results in good estimation performance for any sampling policy.

Table 2. State estimation based on the DC-PF model: the average MSE and its standard deviation (std.) for the GSP-WLS (26) and (27) over different values of $\mu \in [0.01, 1]$, for $q = 48$ and $\sigma^2 = 0.01$.

Measure	Bus Selection Method		
	Random	E-Design	Algorithm 1
Average MSE	0.2479	0.7929	0.0116
std. MSE	0.0116	0.0701	9.45×10^{-6}

Figure 9 demonstrates the performance of the GSP-WLS estimator where the power system operating condition changes throughout the day, using time series data for different values of μ and correspondingly different sampling sets chosen by Algorithm 1. Simulation setup and parameters: For this case study, we use the modified IEEE 118-bus system in a similar manner to [67]. Figure 9a shows the hourly load profile of the system demand, as taken from <http://motor.ece.iit.edu/Data/> (accessed on 1 August 2022). Given a single measure of the loads in all the buses and the total system demand along 24 h, the hourly load at bus n is the relative part of the load (among the other measured loads) from the total system demand. The true value of the state was calculated using the MATPOWER toolbox [68]. It can be seen in Figure 9b that for the MSE for $\mu = 0.01, 0.1, 1$ is approximately the same. These results show that the proposed GSP-WLS estimator is robust to the tuning parameter in this range, even under changing operating conditions. More strategies for choosing μ are described in the literature (e.g., see [61]).

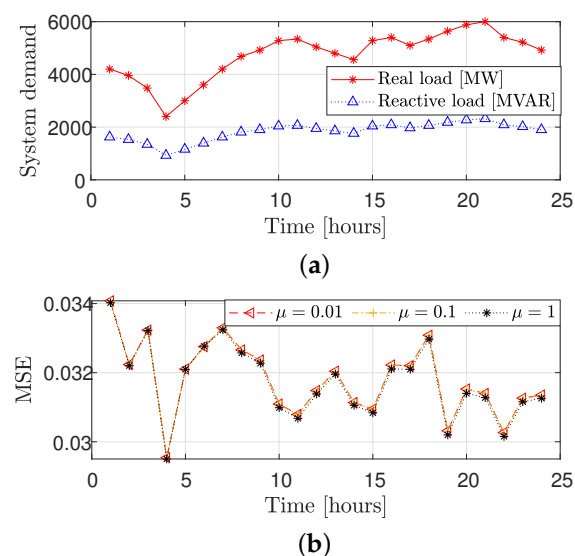


Figure 9. State estimation based on the DC-PF model: the hourly system demand (a) and the corresponding MSE of the GSP-WLS estimator (26) and (27) (b) versus time over 24 h for $q = 48$ buses and $\sigma^2 = 0.01$.

6.3. State Estimation and Sampling under the AC-PF Model

In the following, we discuss the estimation of both the phases and the magnitudes of the states using the estimators:

1. The Gauss–Newton implementation of the pm-WLS estimator from [11] by Algorithm 2 with the appropriate replacements of the regularization terms, i.e., where $\hat{\mathbf{x}}_{prior} = \mathbf{x}_0$ and $\mathbf{R}_{prior}^{-1} = \mathbf{I}$ are used instead of $\bar{\mathbf{L}}(\mu_\theta, \mu_v)$ (see the remarks after (57)).
2. The mc method from [5], implemented using Equations (6), (8), (12a), and (12c) from [5], where the low-rank matrix used in this method, composed of the real and the imaginary parts of \mathbf{v} . In addition, we added to this method the current measurements as inputs (instead of the power flow measurements) for a fair comparison. The implementation was conducted by the SDP solver of CVX [60].
3. The proposed regularized Gauss–Newton method for implementing the GSP-WLS estimator from Algorithm 2, with the regularization parameters $\mu_\theta = 0.045$ and $\mu_v = 10$.

In the Gauss–Newton-based methods, the maximal number of iterations is set to $l = 20$ and $\delta = 10^{-8}$ in Algorithm 2.

In Figure 10a the MSE of phase estimation by the GSP-WLS and pm-WLS estimators (both implemented by the regularized Gauss–Newton method) are presented versus the number of measured buses, q , for the sampling policies (i)–(iii). Similarly, in Figure 10b, the MSE of the magnitude estimation is presented. It can be seen that the MSE decreases as the number of measured buses increases for both the magnitudes and the phases. Moreover, the GSP-WLS estimator outperforms the pm-WLS and the mc estimators for any sampling policy. Figure 10a shows that for each sampling policy, the MSEs of the GSP-WLS, and the pm-WLS estimators for the phases separate from each other when the system observability can no longer be guaranteed (i.e., where $q < 72$ for the random sampling, $q < 76$ for the E-design sampling, and $q < 79$ for Algorithm 1). In particular, where $q = 118$ (i.e., full observability with probability 1), the performances of the estimators coincide. Figure 10b demonstrates that for $q = 118$, the MSEs of the GSP-WLS and the pm-WLS estimators for magnitude are equal. For an observable system with $q < 118$, the MSEs of the estimators coincide for each sampling policy separately. Finally, when the system becomes unobservable (i.e., where $q < 72$ for the random sampling, $q < 76$ for the E-design sampling, and $q < 79$ for Algorithm 1), the MSEs of the estimators split, and the MSE of the GSP-WLS is lower.

It should be noted that the mc estimator implementation by CVX has higher computational complexity and a lot of tuning parameters in comparison to the other methods. Finally, it can be seen that the sampling policy from Algorithm 1 results in a significantly lower MSE than that obtained for the random and the E-design sampling policies for the pm-WLS and the GSP-WLS estimators.

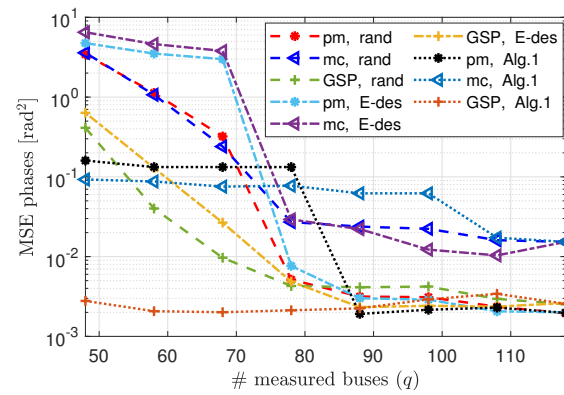
In Figure 11, we examine the influence of the tuning parameters μ_θ and μ_v on the estimation performance of the phases and magnitudes, where the sampling set is chosen by Algorithm 1. It can be seen that for $\mu_\theta, \mu_v \in [0.01, 10]$, the MSE is approximately constant. Hence, choosing any μ_θ and μ_v in this range will obtain a good result.

6.4. Detection of Bad Data

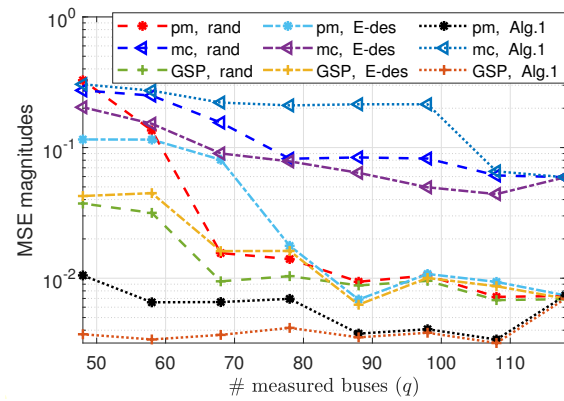
In the following, the application of bad data detection based on the proposed estimator is demonstrated, based on the description in Section 4.4. We compare the performance of the following detectors:

- The LNR test from (34), the $J(\theta)$ test from (35), and the GFT-based detector from [7], after substituting either $\hat{\theta} = \hat{\theta}^{\text{pm-WLS}}$ from (29) or $\hat{\theta} = \hat{\theta}^{\text{GSP-WLS}}$ from (26) and (27). In the GFT-based method, the cutoff frequency is chosen such that in normal states, 35% of the frequencies pass the filter.
- The L_∞ detector from [69] with $\Sigma_x = \mathbf{I}$.

- The Laplacian-Regularized detector (Lap.Reg.) from our previous work in [41], which also uses the GSP-WLS estimator, $\hat{\theta}^{\text{GSP-WLS}}$.



(a)



(b)

Figure 10. State estimation under the AC–PF model: the phase MSE (a) and the magnitude MSE (b) of the pm–WLS [11], mc [5], and GSP–WLS (Algorithm 2) estimators for random, E–design [38], and Algorithm 1 bus selection policies versus the number of buses, q .

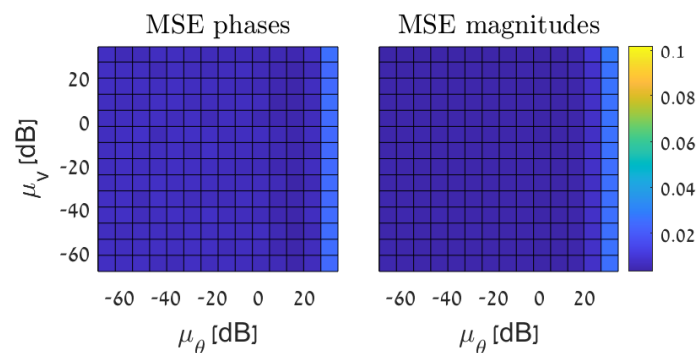


Figure 11. State estimation based on the AC–PF model: the MSE of the GSP–WLS (Algorithm 2) for phase estimation (left) and magnitude estimation (right) with $q = 48$ buses and $\sigma^2 = 0.01$ versus the value of the tuning parameters, μ_θ and μ_v .

Simulation setup and parameters: The detection performance is evaluated for partially observable systems with only power injection measurements, i.e., $\mathbf{L}_{v,S} = \mathbf{H}_{v,S}$, with $q = 48$ buses that are chosen by Algorithm 1. Constant noise ($\Delta z = 10\sigma$) has been added onto three random measurements, i.e., $z_i + \Delta z$ at the i th bus, where z_i is obtained by (21).

In Figure 12, the detection performance of the proposed detectors is demonstrated by the receiver operating characteristic (ROC) curves. This figure shows that the detectors

that are based on the GSP-WLS estimator outperform the detectors that are based on the pm-WLS estimator, as well as the L_∞ detector. These results are aligned with the estimation results, since the MSE of the state estimation by the GSP-WLS estimator is significantly lower than the MSE obtained by the pm-WLS estimator, as shown in Figure 7.

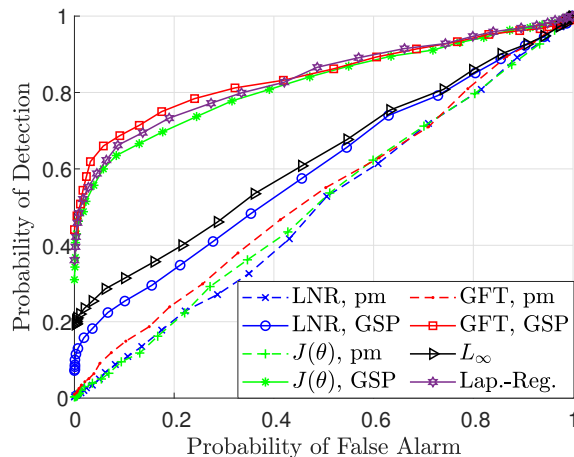


Figure 12. Bad data detection: the ROC of the LNR test (34), $J(\theta)$ test (35), GFT-based detector [7] implemented with the GSP-WLS (26) and (27), and the pm-WLS [11] estimators, L_∞ [69], and the Laplacian-regularized [41] detectors with $q = 48$ buses.

7. Discussion

Utilization and Implications: The proposed GSP-WLS estimator has the potential to significantly improve state estimation in unobservable power systems, which can be caused by various reasons, such as communication issues, cyber attacks, and legal or economic constraints on measurement locations. By utilizing GSP techniques, the proposed estimator can achieve more accurate estimates with fewer measurements, reducing the cost and complexity of data collection. In addition, the proposed sensor placement strategy optimizes the performance of the estimator, further improving its accuracy. These GSP techniques have important implications for the reliability and stability of power systems, as well as for the development of advanced control and optimization algorithms. The research also demonstrates the capabilities of GSP in power system networks, highlighting the potential for its use in other tasks within the field. Overall, this research aims to increase the stability and reliability of electrical grids by developing monitoring techniques. Additionally, this study is closely related to the control and management of renewable energy sources, since using enhanced estimation, sensor allocation, and attack detection methods, the ability of the grid to deal with randomness in the system that stems from the sources increases.

Limitations: This research provides a new GSP framework for state estimation and bad data detection. However, it has some limitations that should be considered in future research. First, it only uses the most recent vector of measurements. While this is a good policy when the system may be under abrupt changes, the performance of the estimator can be improved by incorporating historical measurements from previous time steps. Second, the topology of the network is assumed to be known, which may not always be the case, particularly for distribution systems. In such cases, methods for estimating the topology [29,70] could be used. Third, the Gauss-Newton algorithm used in this study may be computationally expensive for large networks. To address this issue, future research could use low-complexity algorithms to calculate matrix inversion, as described in [1]. Finally, this study has not explored the incorporation of smart meter and PMU measurements into the proposed methods, and this should be considered in future research.

Future research: Future research directions include the incorporation of time series measurements with temporal dependencies to improve the estimation and bad data performance. Additionally, it would be useful to investigate the joint estimation of the system

state and the identification of topology changes. Another potential future direction is the combination of smart meter measurements and PMU data in the proposed method, as mentioned above in the limitations. Finally, a promising area of future research is the integration of deep unfolding or other machine learning and optimization tools to accelerate the regularized Gauss–Newton method. One particularly important area of future research is the development of new graph neural network (GNN) approaches that take advantage of the graph properties of the states that were validated in this study. GNN methods could potentially offer significant improvements in the accuracy and efficiency of state estimation in power systems, depending on the availability of relevant data.

8. Conclusions

This paper proposes a GSP framework for state estimation, sensor allocation, and bad data detection in unobservable power systems. First, the graph smoothness of the phases and the magnitudes of the voltages with respect to the admittance matrix have been validated. Then, the GSP-WLS estimator of the system states has been derived under both the DC-PF and AC-PF models. The GSP-WLS estimator uses the graph smoothness of the state signals as a regularization term, and thus, does not require full observability of the system. It is analytically shown that when the system is observable, the proposed GSP-WLS estimator coincides with the WLS estimator. A greedy algorithm has been introduced to tackle the problem of selecting the sampling set that optimizes the state estimation performance. In addition, it is shown that the GSP-WLS estimator is useful for bad data detection by plugging the estimator into existing detectors. Simulation results demonstrate the potential of the GSP methods in power systems for cases where the system observability is not guaranteed. It is shown that the proposed methods can accurately estimate voltage phases and magnitudes, and detect bad data under partial observability conditions where standard methods cannot, or exhibit poor performance. These results show that advanced sensing and measurement technologies using GSP tools can transform data into useful information and enhance various aspects of power system management.

Author Contributions: The authors' contributions are as follows: conceptualization, L.D., A.K. and T.R.; methodology, L.D., A.K. and T.R.; software, L.D.; validation, L.D. and T.R.; formal analysis, L.D. and T.R.; investigation, L.D.; resources, T.R.; data curation, L.D.; writing—original draft preparation, L.D.; writing—review and editing, T.R.; visualization, L.D.; supervision, T.R.; project administration, T.R.; funding acquisition, L.D. and T.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Israel Ministry of National Infrastructure, Energy, and Water Resources and by the BGU Cyber Security Research Center. L. Dabush is a fellow of the Ariane de Rothschild Women's Doctoral Program.

Data Availability Statement: The IEEE test case systems data in the simulations has been taken from the "Power Systems Test Case Archive", ref. [56], <http://www.ee.washington.edu/research/pstca/> (accessed on 1 June 2019).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

GSP	Graph signal processing
WLS	Weighted least squares
PF	Power flow
DC	Direct current
AC	Alternating current
PSSE	Power systems state estimation
EMS	Energy management system

GFT	Graph Fourier transform
DSP	Digital signal processing
KKT	Karush–Kuhn–Tucker
pm	Pseudo-measurements
MSE	Mean squared error
CRB	Cramér–Rao bound
LNR	Largest normalized residual
FDI	False data injection
ROC	Receiver operating characteristic
GNN	Graph neural network

References

- Abur, A.; Gomez-Exposito, A. *Power System State Estimation: Theory and Implementation*; Dekker, M., Ed.; CRC Press: Boca Raton, FL, USA, 2004.
- Monticelli, A. In *State Estimation in Electric Power Systems: A Generalized Approach*; Springer: Boston, MA, USA, 1999; pp. 39–61, 91–101, 161–199. [\[CrossRef\]](#)
- Mestav, K.R.; Luengo-Rozas, J.; Tong, L. Bayesian State Estimation for Unobservable Distribution Systems via Deep Learning. *IEEE Trans. Power Syst.* **2019**, *34*, 4910–4920. [\[CrossRef\]](#)
- De Almeida, M.C.; Garcia, A.V.; Asada, E.N. Regularized Least Squares Power System State Estimation. *IEEE Trans. Power Syst.* **2012**, *27*, 290–297. [\[CrossRef\]](#)
- Donti, P.L.; Liu, Y.; Schmitt, A.J.; Bernstein, A.; Yang, R.; Zhang, Y. Matrix Completion for Low-Observability Voltage Estimation. *IEEE Trans. Smart Grid* **2020**, *11*, 2520–2530. [\[CrossRef\]](#)
- Kosut, O.; Jia, L.; Thomas, R.J.; Tong, L. Malicious Data Attacks on the Smart Grid. *IEEE Trans. Smart Grid* **2011**, *2*, 645–658. [\[CrossRef\]](#)
- Drayer, E.; Routtenberg, T. Detection of False Data Injection Attacks in Smart Grids Based on Graph Signal Processing. *IEEE Syst. J.* **2020**, *14*, 1886–1896. [\[CrossRef\]](#)
- Morgenstern, G.; Routtenberg, T. Structural-constrained Methods for the Identification of Unobservable False Data Injection Attacks in Power Systems. *IEEE Access* **2022**, *10*, 94169–94185. [\[CrossRef\]](#)
- Buldyrev, S.V.; Parshani, R.; Paul, G.; Stanley, H.E.; Havlin, S. Catastrophic cascade of failures in interdependent networks. *Nature* **2010**, *464*, 1025–1028. [\[CrossRef\]](#)
- Gao, P.; Wang, M.; Ghiocel, S.G.; Chow, J.H.; Fardanesh, B.; Stefopoulos, G. Missing Data Recovery by Exploiting Low-Dimensionality in Power System Synchrophasor Measurements. *IEEE Trans. Power Syst.* **2016**, *31*, 1006–1013. [\[CrossRef\]](#)
- Do Coutto Filho, M.B.; Stacchini de Souza, J.C. Forecasting-Aided State Estimation—Part I: Panorama. *IEEE Trans. Power Syst.* **2009**, *24*, 1667–1677. [\[CrossRef\]](#)
- Clements, K.A. The impact of pseudo-measurements on state estimator accuracy. In Proceedings of the IEEE PES General Meeting, Detroit, MI, USA, 24–28 July 2011; pp. 1–4.
- Zhao, J.; Zhang, G.; Dong, Z.Y.; La Scala, M. Robust Forecasting Aided Power System State Estimation Considering State Correlations. *IEEE Trans. Smart Grid* **2018**, *9*, 2658–2666. [\[CrossRef\]](#)
- Zhao, J.; Gómez-Expósito, A.; Netto, M.; Mili, L.; Abur, A.; Terzija, V.; Kamwa, I.; Pal, B.; Singh, A.K.; Qi, J.; et al. Power System Dynamic State Estimation: Motivations, Definitions, Methodologies, and Future Work. *IEEE Trans. Power Syst.* **2019**, *34*, 3188–3198. [\[CrossRef\]](#)
- Bhela, S.; Kekatos, V.; Veeramachaneni, S. Enhancing Observability in Distribution Grids Using Smart Meter Data. *IEEE Trans. Smart Grid* **2018**, *9*, 5953–5961. [\[CrossRef\]](#)
- Xu, B.; Abur, A. Observability Analysis and Measurement Placement for Systems with PMUs. In Proceedings of the IEEE PES Power Systems Conference and Exposition, New York, NY, USA, 10–13 October 2004; Volume 2, pp. 943–946. [\[CrossRef\]](#)
- Zamzam, A.S.; Liu, Y.; Bernstein, A. Model-Free State Estimation Using Low-Rank Canonical Polyadic Decomposition. *IEEE Contr. Syst. Lett.* **2020**, *5*, 605–610. [\[CrossRef\]](#)
- Paramo, G.; Bretas, A.; Meyn, S. Research Trends and Applications of PMUs. *Energies* **2022**, *15*, 5329. [\[CrossRef\]](#)
- Mestav, K.R.; Tong, L. Learning the Unobservable: High-Resolution State Estimation via Deep Learning. In Proceedings of the 2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton), Monticello, IL, USA, 24–27 September 2019; pp. 171–176. [\[CrossRef\]](#)
- Zhang, J.; Welch, G.; Bishop, G.; Huang, Z. Reduced Measurement-space Dynamic State Estimation (ReMeDySE) for power systems. In Proceedings of the IEEE Trondheim PowerTech, Trondheim, Norway, 19–23 June 2011; pp. 1–7. [\[CrossRef\]](#)
- Wang, W.; Yu, N.; Rahmatian, F.; Pandey, S. Where to Install Distribution Phasor Measurement Units to Obtain Accurate State Estimation Results? In Proceedings of the 2022 IEEE Power & Energy Society General Meeting (PESGM), Denver, CO, USA, 17–21 July 2022; pp. 1–5. [\[CrossRef\]](#)
- Zamzam, A.S.; Sidiropoulos, N.D. Physics-Aware Neural Networks for Distribution System State Estimation. *IEEE Trans. Power Syst.* **2020**, *35*, 4347–4356. [\[CrossRef\]](#)

23. Huang, M.; Wei, Z.; Lin, Y. Forecasting-aided State Estimation Based on Deep Learning for Hybrid AC/DC Distribution Systems. *Appl. Energy* **2022**, *306*, 118119. [[CrossRef](#)]
24. Do Coutto Filho, M.B.; de Souza, J.C.S.; Schilling, M.T. Generating High Quality Pseudo-Measurements to Keep State Estimation Capabilities. In Proceedings of the IEEE Lausanne Power Tech, Lausanne, Switzerland, 1–5 July 2007; pp. 1829–1834. [[CrossRef](#)]
25. Yaniv, A.; Kumar, P.; Beck, Y. Towards adoption of GNNs for power flow applications in distribution systems. *Electr. Power Syst. Res.* **2023**, *216*, 109005. [[CrossRef](#)]
26. Pourali, M.; Mosleh, A. A Functional Sensor Placement Optimization Method for Power Systems Health Monitoring. *IEEE Trans. Ind. Appl.* **2013**, *49*, 1711–1719. [[CrossRef](#)]
27. Soltan, S.; Mazauric, D.; Zussman, G. Analysis of Failures in Power Grids. *IEEE Control Netw. Syst.* **2017**, *4*, 288–300. [[CrossRef](#)]
28. Deka, D.; Backhaus, S.; Chertkov, M. Estimating Distribution Grid Topologies: A Graphical Learning Based Approach. In Proceedings of the PSCC, Genoa, Italy, 20–24 June 2016.
29. Grotas, S.; Yakoby, Y.; Gera, I.; Routtenberg, T. Power Systems Topology and State Estimation by Graph Blind Source Separation. *IEEE Trans. Signal Process.* **2019**, *67*, 2036–2051. [[CrossRef](#)]
30. Weng, Y.; Negi, R.; Ilić, M.D. Graphical model for state estimation in electric power systems. In Proceedings of the IEEE International Conference on Smart Grid Communications (SmartGridComm), Vancouver, BC, Canada, 21–24 October 2013; pp. 103–108.
31. Giannakis, G.B.; Kekatos, V.; Gatsis, N.; Kim, S.J.; Zhu, H.; Wollenberg, B.F. Monitoring and Optimization for Power Grids: A Signal Processing Perspective. *IEEE Signal Process. Mag.* **2013**, *30*, 107–128. [[CrossRef](#)]
32. Ishizaki, T.; Chakraborty, A.; Imura, J. Graph-Theoretic Analysis of Power Systems. *Proc. IEEE* **2018**, *106*, 931–952. [[CrossRef](#)]
33. Guo, L.; Zhao, C.; Low, S.H. Graph Laplacian Spectrum and Primary Frequency Regulation. In Proceedings of the 2018 IEEE Conference on Decision and Control (CDC), Miami, FL, USA, 17–19 December 2018.
34. Dvijotham, K.; Van Hentenryck, P.; Chertkov, M.; Misra, S.; Vuffray, M. Graphical Models for Optimal Power Flow. *Constraints* **2016**, *22*, 24–49. [[CrossRef](#)]
35. Retière, N.; Ha, D.T.; Caputo, J.G. Spectral Graph Analysis of the Geometry of Power Flows in Transmission Networks. *IEEE Syst. J.* **2020**, *14*, 2736–2747. [[CrossRef](#)]
36. Sandryhaila, A.; Moura, J.M.F. Discrete Signal Processing on Graphs. *IEEE Trans. Signal Process.* **2013**, *61*, 1644–1656. [[CrossRef](#)]
37. Shuman, D.I.; Narang, S.K.; Frossard, P.; Ortega, A.; Vandergheynst, P. The Emerging Field of Signal Processing on Graphs: Extending High-dimensional Data Analysis to Networks and Other Irregular Domains. *IEEE Signal Process. Mag.* **2013**, *30*, 83–98. [[CrossRef](#)]
38. Chen, S.; Varma, R.; Sandryhaila, A.; Kovačević, J. Discrete Signal Processing on Graphs: Sampling Theory. *IEEE Trans. Signal Process.* **2015**, *63*, 6510–6523. [[CrossRef](#)]
39. Marques, A.G.; Segarra, S.; Leus, G.; Ribeiro, A. Sampling of Graph Signals with Successive Local Aggregations. *IEEE Trans. Signal Process.* **2016**, *64*, 1832–1843. [[CrossRef](#)]
40. Ramakrishna, R.; Scaglione, A. Grid-Graph Signal Processing (Grid-GSP): A Graph Signal Processing Framework for the Power Grid. *IEEE Trans. Signal Process.* **2021**, *69*, 2725–2739. [[CrossRef](#)]
41. Dabush, L.; Routtenberg, T. Detection of False Data Injection Attacks in Unobservable Power Systems by Laplacian Regularization. In Proceedings of the 2022 IEEE 12th Sensor Array and Multichannel Signal Processing Workshop (SAM), Trondheim, Norway, 20–23 June 2022.
42. Hasnat, M.A.; Rahnamay-Naeini, M. A Graph Signal Processing Framework for Detecting and Locating Cyber and Physical Stresses in Smart Grids. *IEEE Trans. Smart Grid* **2022**, *13*, 3688–3699. [[CrossRef](#)]
43. Saha, S.S.; Scaglione, A.; Ramakrishna, R.; Johnson, N.G. Distribution Systems AC State Estimation via Sparse AMI Data Using Graph Signal Processing. *IEEE Trans. Smart Grid* **2022**, *13*, 3636–3649. [[CrossRef](#)]
44. Zheng, M.; Bu, J.; Chen, C.; Wang, C.; Zhang, L.; Qiu, G.; Cai, D. Graph Regularized Sparse Coding for Image Representation. *IEEE Trans. Image Process.* **2011**, *20*, 1327–1336. [[CrossRef](#)] [[PubMed](#)]
45. Elmoataz, A.; Lezoray, O.; Boudgoux, S. Nonlocal Discrete Regularization on Weighted Graphs: A Framework for Image and Manifold Processing. *IEEE Trans. Image Process.* **2008**, *17*, 1047–1060. [[CrossRef](#)] [[PubMed](#)]
46. Shahid, N.; Kalofolias, V.; Bresson, X.; Bronstein, M.; Vandergheynst, P. Robust principal component analysis on graphs. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 2812–2820.
47. Wang, F.; Zhang, C. Label propagation through linear neighborhoods. *IEEE Trans. Knowl. Data Eng.* **2007**, *20*, 55–67. [[CrossRef](#)]
48. Belkin, M.; Matveeva, I.; Niyogi, P. Regularization and semi-supervised learning on large graphs. In *International Conference on Computational Learning Theory*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 624–638.
49. Chapelle, O.; Scholkopf, B.; Zien, A. *Semi-Supervised Learning*; The MIT Press: Cambridge, MA, USA, 2006.
50. Jia, J.; Schaub, M.T.; Segarra, S.; Benson, A.R. Graph-Based Semi-Supervised and Active Learning for Edge Flows. In Proceedings of the International Conference on Knowledge Discovery and Data Mining, Anchorage, AK, USA, 4–8 August 2019; pp. 761–771.
51. Ortega, A.; Frossard, P.; Kovačević, J.; Moura, J.M.F.; Vandergheynst, P. Graph Signal Processing: Overview, Challenges, and Applications. *Proc. IEEE* **2018**, *106*, 808–828. [[CrossRef](#)]
52. Ramakrishna, R.; Wai, H.T.; Scaglione, A. A User Guide to Low-Pass Graph Signal Processing and Its Applications: Tools and Applications. *IEEE Signal Process. Mag.* **2020**, *37*, 74–85. [[CrossRef](#)]
53. Newman, M. *Networks: An Introduction*; Oxford University Press, Inc.: New York, NY, USA, 2010.

54. Aien, M.; Fotuhi-Firuzabad, M.; Aminifar, F. Probabilistic Load Flow in Correlated Uncertain Environment Using Unscented Transformation. *IEEE Trans. Power Syst.* **2012**, *27*, 2233–2241. [[CrossRef](#)]
55. Chakhchoukh, Y.; Panciatici, P.; Mili, L. Electric Load Forecasting Based on Statistical Robust Methods. *IEEE Trans. Power Syst.* **2011**, *26*, 982–991. [[CrossRef](#)]
56. Christie, R.D. Power Systems Test Case Archive. 1999. Available online: <http://www.ee.washington.edu/research/pstca/> (accessed on 1 June 2019). [[CrossRef](#)]
57. Puy, G.; Pérez, P. Structured sampling and fast reconstruction of smooth graph signals. *Inf. Inference J. IMA* **2017**, *26*, 1770–1785. [[CrossRef](#)]
58. Pang, J.; Cheung, G. Graph Laplacian Regularization for Image Denoising: Analysis in the Continuous Domain. *IEEE Trans. Image Process.* **2018**, *7*, 657–688. [[CrossRef](#)]
59. Kalofolias, V. How to learn a graph from smooth signals. In Proceedings of the 19th International Conference on Artificial Intelligence and Statistics, Cadiz, Spain, 9–11 May 2016; pp. 920–929.
60. Boyd, S.; Vandenberghe, L. *Convex Optimization*; Cambridge University Press: New York, NY, USA, 2004.
61. van Wieringen, W.N. Lecture Notes on Ridge Regression. *arXiv* **2015**, arXiv:1509.09169.
62. Madani, R.; Lavaei, J.; Baldick, R. Convexification of Power Flow Equations in the Presence of Noisy Measurements. *IEEE Trans. Autom. Control* **2019**, *64*, 3101–3116. [[CrossRef](#)]
63. Donmez, B.; Abur, A. Sparse Estimation Based External System Line Outage Detection. In Proceedings of the PSCC, Genoa, Italy, 20–24 June 2016; pp. 1–6. [[CrossRef](#)]
64. Zhao, Y.; Chen, J.; Goldsmith, A.; Poor, H.V. Identification of Outages in Power Systems with Uncertain States and Optimal Sensor Locations. *IEEE J. Sel. Top. Signal Process.* **2014**, *8*, 1140–1153. [[CrossRef](#)]
65. Kay, S.M. *Fundamentals of Statistical Signal Processing: Estimation Theory*; Prentice Hall PTR: Englewood Cliffs, NJ, USA, 1993; Volume 1.
66. Djuric, P.M.; Richard, C. Sampling and Recovery of Graph Signals. In *Cooperative and Graph Signal Processing: Principles and Applications*; Academic Press: Cambridge, MA, USA, 2018; Chapter 9.
67. Wu, H.; Shahidehpour, M.; Khodayar, M.E. Hourly Demand Response in Day-Ahead Scheduling Considering Generating Unit Ramping Cost. *IEEE Trans. Power Syst.* **2013**, *28*, 2446–2454. [[CrossRef](#)]
68. Zimmerman, R.D.; Murillo-Sánchez, C.E.; Thomas, R.J. MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education. *IEEE Trans. Power Syst.* **2011**, *26*, 12–19. [[CrossRef](#)]
69. Kosut, O.; Jia, L.; Thomas, R.J.; Tong, L. Limiting False Data Attacks on Power System State Estimation. In Proceedings of the 2010 44th Annual Conference on Information Sciences and Systems (CISS), Princeton, NJ, USA, 17–19 March 2010; pp. 1–6. [[CrossRef](#)]
70. Halihal, M.; Routtenberg, T. Estimation of the Admittance Matrix in Power Systems Under Laplacian and Physical Constraints. In Proceedings of the ICASSP 2022—2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Singapore, 22–27 May 2022; pp. 5972–5976. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.