OPTIMAL SENSOR PLACEMENT FOR HYBRID STATE ESTIMATION IN SMART GRID

Xiao Li, Anna Scaglione and Tsung-Hui Chang

Dept. of Electrical and Computer Engineering, University of California, Davis, CA, 95616, USA

ABSTRACT

A critical task in smart grid is to gain situational awareness by performing state estimation. In this paper, we consider the problem of placing a type of special sensors, called Phasor Measurement Units (PMU), to optimize the performance and convergence of state estimation. We derive a metric to evaluate how the placement impacts the *convergence* and *accuracy* of state estimation solved by Gauss-Newton (GN) algorithm. Using the proposed metric, we formulate and solve the placement problem as a semi-definite program (SDP). Simulations of the IEEE 30 and 118 systems corroborate our analysis, showing that the proposed placement stabilizes and accelerates state estimation, while maintaining optimal estimation performance.

Index Terms— Optimal placement, convergence, estimation

1. INTRODUCTION

Power system state estimation (PSSE) using non-linear measurements from the Supervisory Control and Data Acquisition (SCADA) systems is plagued by numerical issues. With the GPS technology, a new type of sensors called Phasor Measurement Units (PMU) deployed in the Wide-Area Measurement System (WAMS) can nowadays provide synchronized voltage and current phasor readings at each instrumented bus (i.e., substation), benefiting greatly state estimation [1] because it becomes a inear least squares problem [2].

Since PMU devices are expensive, their placements are strategically optimized to reap the greatest benefits in terms of *observability* [3] under a certain budget [4, 5]. When the system is observable, the state becomes uniquely identifiable from the corresponding measurements [6]. There is vast literature on minimizing the deployment cost under the observability constraint (see e.g. [7–17]). These works are usually formulated as integer programs with different cost functions and solved via a variety of numerical techniques.

The PMU placement can also be optimized with respect to estimation accuracy [18–21]. For example, [19] uses a two-stage approach that first guarantees observability and then refines the placement for estimation performance. In [22, 23], instead, PMUs are placed iteratively on buses with the largest errors (individual or sum), until a cost budget is met. A greedy method was proposed in [20] for PMU placement by minimizing the estimation errors of the augmented PSSE in polar coordinates using voltage and linearized power injection measurements. The same problem is solved via convex relaxation in [21] using Cartesian coordinates for PMU data. The mutual information between sensor measurements and state vector was mentioned in [24] as a unified metric of observability and accuracy.

In light of the increasing interest in using hybrid measurements from both PMU and SCADA systems, we revisit the problem of optimizing the PMU placement in order to account also for the convergence of the iterative algorithm in state estimation when SCADA data are used [25]. The question we try to answer in this paper is how the PMU placement affects the stability and rate of convergence of PSSE, and whether a judicious placement can further stabilize and accelerate state estimation while enhancing its accuracy.

The contribution of this paper is the derivation of the Joint Accuracy and Convergence (JAC) metric to evaluate the performance and robustness of the hybrid PSSE for a given sensor placement. We also optimize the PMU placement with respect to the JAC metric via semidefinite programming (SDP). Finally, we numerically show the performance of the proposed placement with alternative designs.

2. POWER SYSTEM STATE ESTIMATION

We consider a power grid with N *buses* (i.e., substations), representing interconnections, generators or loads. They are denoted by the set $\mathcal{N} \triangleq \{1, \cdots, N\}$, which form the edge set $\mathcal{E} \triangleq \{(n,m)\}$ with cardinality $|\mathcal{E}| = L$, with $\{(n,m)\}$ denoting the transmission line between n and m. Furthermore, we define $\mathcal{N}(n) \triangleq \{m : (n,m) \in \mathcal{E}\}$ as the neighbor of bus n and let $L_n = |\mathcal{N}(n)|$. The Energy Management Systems (EMS) at control centers collect measurements on certain buses and transmission lines to estimate the state of the power system, i.e., the voltage phasor $V_n \in \mathbb{C}$ at each bus $n \in \mathcal{N}$. In this paper, we consider the Cartesian coordinate representation using the real and imaginary components of the complex voltage phasors $\mathbf{v} = [\Re\{V_1\}, \cdots, \Re\{V_N\}, \Im\{V_1\}, \cdots, \Im\{V_N\}]^T$.

2.1. Measurement Model and State Estimation

Given that there are 2 complex injection measurements at each bus, and 4 complex flow measurements associated with each line, which amount to twice as many real variables, the ensemble of all measurements is of length $M = 4N + 8L$ and represented by an aggregate vector partitioned into four sections $\mathbf{z} = [\mathbf{z}_{\mathcal{V}}^T, \mathbf{z}_{\mathcal{C}}^T, \mathbf{z}_{\mathcal{I}}^T, \mathbf{z}_{\mathcal{F}}^T]^T$, containing the length-$2N$ voltage phasor $\mathbf{z}_{\mathcal{V}}$ and power injection vector $\mathbf{z}_{\mathcal{I}}$ at bus $n \in \mathcal{N}$, the length-$4L$ current phasor $\mathbf{z}_{\mathcal{C}}$ and power flow vector $\mathbf{z}_{\mathcal{F}}$ on line $(n,m) \in \mathcal{E}$ at bus n. Defining the power flow equations $\mathbf{f}_{(\cdot)}(\mathbf{v})$ in Appendix A and letting $\bar{\mathbf{v}}$ be the true system state, the individual set $\mathbf{z}_{(\cdot)} = \mathbf{f}_{(\cdot)}(\bar{\mathbf{v}}) + \mathbf{r}_{(\cdot)}$ contains observations corrupted by measurement noise $\mathbf{r}_{(\cdot)}$ that arises from instrumentation imprecision with zero mean and a covariance matrix $\mathbf{R} \triangleq \mathbb{E}\{\mathbf{r}\mathbf{r}^T\}$. The combined noisy measurement model is

$$\mathbf{z} = \mathbf{f}(\bar{\mathbf{v}}) + \mathbf{r}, \tag{1}$$

where $\mathbf{r} = [\mathbf{r}_{\mathcal{V}}^T, \mathbf{r}_{\mathcal{C}}^T, \mathbf{r}_{\mathcal{I}}^T, \mathbf{r}_{\mathcal{F}}^T]^T$ is the aggregate noise and $\mathbf{f}(\mathbf{v}) = [\mathbf{f}_{\mathcal{V}}^T(\mathbf{v}), \mathbf{f}_{\mathcal{C}}^T(\mathbf{v}), \mathbf{f}_{\mathcal{I}}^T(\mathbf{v}), \mathbf{f}_{\mathcal{F}}^T(\mathbf{v})]^T$. In practice, the collected observations are a subset of $\mathbf{z}$ in (1). For convenience, we introduce an appropriate $M \times M$) diagonal mask $\mathbf{J}$ having 1 on its diagonal if and only if that measurement is collected, giving

$$\mathbf{c} = \mathbf{J}\mathbf{f}(\bar{\mathbf{v}}) + \mathbf{J}\mathbf{r}. \tag{2}$$

Assuming $\mathbf{R} = \text{diag}\,[\mathbf{R}_{\mathcal{V}}, \mathbf{R}_{\mathcal{I}}, \mathbf{R}_{\mathcal{I}}, \mathbf{R}_{\mathcal{F}}]$ with $\mathbf{R}_{(\cdot)} = \sigma_{(\cdot)}^2\mathbf{I}$ for some $\sigma_{\mathcal{V}}$, $\sigma_{\mathcal{C}}$, $\sigma_{\mathcal{I}}$ and $\sigma_{\mathcal{F}}$, the state is then estimated as [26, 27]

$$\widehat{\mathbf{v}} = \arg\min_{\mathbf{v}\in\mathbb{V}} \ \|\tilde{\mathbf{c}} - \tilde{\mathbf{f}}(\mathbf{v})\|^2, \tag{3}$$

where $\tilde{\mathbf{c}} = \mathbf{R}^{-\frac{1}{2}}\mathbf{c}$ and $\tilde{\mathbf{f}}(\mathbf{v}) = \mathbf{R}^{-\frac{1}{2}}\mathbf{J}\mathbf{f}(\mathbf{v})$ are the pre-whitened counterparts of $\mathbf{c}$ and $\mathbf{f}(\mathbf{v})$, and $\mathbb{V}$ is the state space. For discussions,

we let $\tilde{\mathbf{c}} \triangleq [\tilde{\mathbf{c}}_{\mathcal{V}}^T, \tilde{\mathbf{c}}_{\mathcal{C}}^T, \tilde{\mathbf{c}}_{\mathcal{I}}^T, \tilde{\mathbf{c}}_{\mathcal{F}}^T]^T$ and $\mathbf{J} \triangleq \mathrm{diag}[\mathbf{J}_{\mathcal{V}}, \mathbf{J}_{\mathcal{C}}, \mathbf{J}_{\mathcal{I}}, \mathbf{J}_{\mathcal{F}}]$ where $\mathbf{J}_{\mathcal{V}}$, $\mathbf{J}_{\mathcal{C}}$, $\mathbf{J}_{\mathcal{I}}$ and $\mathbf{J}_{\mathcal{F}}$ are the masks for each type of measurement. The Jacobian $\tilde{\mathbf{F}}(\mathbf{v}) = \mathbf{R}^{-\frac{1}{2}}\mathbf{J}\mathrm{d}\mathbf{f}(\mathbf{v})/\mathrm{d}\mathbf{v}^T = \mathbf{R}^{-\frac{1}{2}}\mathbf{J}\mathbf{F}(\mathbf{v})$ can be computed from $\mathbf{F}(\mathbf{v}) \triangleq \mathrm{d}\mathbf{f}(\mathbf{v})/\mathrm{d}\mathbf{v}^T$ given in Appendix A.

2.2. Estimation Performance using Gauss-Newton Algorithm

The Gauss-Newton (GN) algorithm is typically used to solve (3)

$$\mathbf{v}^{k+1} = \mathbf{v}^k + \mathbf{d}^k, \quad k = 1, 2, \cdots \tag{4}$$

with an initializer $\mathbf{v}^0$ and the iterative descent

$$\mathbf{d}^k = \left[\tilde{\mathbf{F}}^T(\mathbf{v}^k)\tilde{\mathbf{F}}(\mathbf{v}^k)\right]^{-1}\tilde{\mathbf{F}}^T(\mathbf{v}^k)\left[\tilde{\mathbf{c}} - \tilde{\mathbf{f}}(\mathbf{v}^k)\right]. \tag{5}$$

Because PMUs directly measure the state, it is natural to exploit them as a good initializer. Here, we propose to choose the initializer $\mathbf{v}^0$ matching PMU measurements wherever available, with the rest provided by an *arbitrary initializer* $\mathbf{s}_{\mathcal{V}}$ (e.g., a stale or nominal estimate). We define the PMU placement vector $\boldsymbol{\mathcal{V}} \triangleq [\mathcal{V}_1, \cdots, \mathcal{V}_N]^T$ with $\mathcal{V}_n \in \{0, 1\}$ indicating whether the n-th bus has a PMU and $\mathbf{J}_{\mathcal{V}} = \mathbf{I}_2 \otimes \mathrm{diag}(\boldsymbol{\mathcal{V}})$. The initializer is then expressed as

$$\mathbf{v}^0(\boldsymbol{\mathcal{V}}) = \mathbf{J}_{\mathcal{V}}\mathbf{z}_{\mathcal{V}} + (\mathbf{I}_{2N} - \mathbf{J}_{\mathcal{V}})\mathbf{s}_{\mathcal{V}}. \tag{6}$$

Due to the non-convex nature of the problem, there are multiple fixed points $\mathbf{v}^\star$ of the update in (4) satisfying the first order condition

$$\tilde{\mathbf{F}}^T(\mathbf{v}^\star)\Big(\tilde{\mathbf{c}} - \tilde{\mathbf{f}}(\mathbf{v}^\star)\Big) = \mathbf{0}, \tag{7}$$

where the estimate $\widehat{\mathbf{v}}$ in (3) corresponds to one of them. Thus, the convergence of the iterate $\mathbf{v}^k$ to $\widehat{\mathbf{v}}$ has been a critical issue in PSSE because it is sensitive to the initializer $\mathbf{v}^0(\boldsymbol{\mathcal{V}})$. If the iterate converges to the estimate $\lim_{k\to\infty}\mathbf{v}^k = \widehat{\mathbf{v}}$, the estimation error is bounded as $\lim_{k\to\infty}\left\|\mathbf{v}^{k+1} - \bar{\mathbf{v}}\right\| \le \|\widehat{\mathbf{v}} - \bar{\mathbf{v}}\|$. If $\mathbf{r}$ in (1) is Gaussian, the solution $\widehat{\mathbf{v}}$ in (3) is the Maximum Likelihood (ML) estimate, whose error covariance is bounded by the Cramér-Rao Bound (CRB)

$$\mathbb{E}\left[\|\widehat{\mathbf{v}} - \bar{\mathbf{v}}\|^2\right] \ge \mathrm{Tr}\left[(\tilde{\mathbf{F}}^T(\bar{\mathbf{v}})\tilde{\mathbf{F}}(\bar{\mathbf{v}}))^{-1}\right], \tag{8}$$

where $\tilde{\mathbf{F}}^T(\bar{\mathbf{v}})\tilde{\mathbf{F}}(\bar{\mathbf{v}})$ is the Fisher Information Matrix (FIM), which depends on the measurement model and the state of the system $\bar{\mathbf{v}}$.

3. AN INTRINSIC METRIC FOR PMU PLACEMENT

Given that the selection $\mathbf{J}$ affects the algorithm convergence and estimation performance, a typical criterion is to guarantee system observability [6]. Here we assume that SCADA measurements are given and thus $\mathbf{J}_{\mathcal{I}}, \mathbf{J}_{\mathcal{F}}$ are known, and focus on the design of PMU placement $\mathbf{J}_{\mathcal{V}}, \mathbf{J}_{\mathcal{C}}$. Thus, we are particularly concerned with capturing the convergence and accuracy in an appropriate metric that also reflects the system observability, and further use that metric to optimize our PMU placement by seeking a deployment that jointly lowers the bound in (8), excludes solutions that make the state unobservable, and stabilizes and accelerates the algorithm convergence.

3.1. System Observability

Observability analysis is typically performed using the load flow model [28], and recently the PMU model [7–17, 21]. The usual way [6] is to examine the invertibility of the Jacobian $\tilde{\mathbf{F}}(\mathbf{v})$, which requires the GN Hessian matrix to be invertible for all $\mathbf{v} \in \mathbb{V}$, i.e.,

$$\beta = \inf_{\mathbf{v}\in\mathbb{V}} \lambda_{\min}\left[\tilde{\mathbf{F}}^T(\mathbf{v})\tilde{\mathbf{F}}(\mathbf{v})\right] > 0, \tag{9}$$

where $\lambda_{\min}[\cdot]$ represents the minimum eigenvalue. Obviously this observability metric depends on the PMU placement $\mathbf{J}_{\mathcal{V}}$ and $\mathbf{J}_{\mathcal{C}}$. If each installed PMU captures the voltage and all incident current measurements on that bus as in [21], $\mathbf{J}_{\mathcal{C}}$ is determined by $\mathbf{J}_{\mathcal{V}}$. Then finally, the GN Hessian is decomposed into two components[1]

$$\tilde{\mathbf{F}}^T(\mathbf{v})\tilde{\mathbf{F}}(\mathbf{v}) = \mathcal{P}(\boldsymbol{\mathcal{V}}) + \mathcal{S}(\mathbf{v}\mathbf{v}^T), \tag{10}$$

where $\mathcal{P}(\cdot) : \mathbb{R}^{2N} \to \mathbb{R}^{2N\times 2N}$ depends on the PMU placement $\boldsymbol{\mathcal{V}}$

$$\mathcal{P}(\boldsymbol{\mathcal{V}}) = \sum_{n=1}^{N}\mathcal{V}_n\left(\frac{\mathbf{I}_2 \otimes \mathbf{e}_n\mathbf{e}_n^T}{\sigma_{\mathcal{V}}^2} + \frac{\mathbf{H}_{I,n}^T\mathbf{H}_{I,n} + \mathbf{H}_{J,n}^T\mathbf{H}_{J,n}}{\sigma_{\mathcal{C}}^2}\right),$$

and $\mathcal{S}(\cdot) : \mathbb{R}^{2N\times 2N} \to \mathbb{R}^{2N\times 2N}$ depends on the SCADA data

$$\mathcal{S}(\mathbf{V}) = \mathbf{H}_{\mathcal{I}}^T\left(\mathbf{J}_{\mathcal{I}}\mathbf{R}_{\mathcal{I}}\mathbf{J}_{\mathcal{I}}^T \otimes \mathbf{V}\right)\mathbf{H}_{\mathcal{I}} + \mathbf{H}_{\mathcal{F}}^T\left(\mathbf{J}_{\mathcal{F}}\mathbf{R}_{\mathcal{F}}\mathbf{J}_{\mathcal{F}}^T \otimes \mathbf{V}\right)\mathbf{H}_{\mathcal{F}},$$

where $\mathbf{H}_{\mathcal{I}}$, $\mathbf{H}_{\mathcal{F}}$, $\mathbf{H}_{I,n}$, $\mathbf{H}_{J,n}$ are defined in (23) in Appendix A.

Remark 1. *Note that the observability metric β can also be an accuracy metric* $\mathrm{Tr}\left[(\tilde{\mathbf{F}}^T(\bar{\mathbf{v}})\tilde{\mathbf{F}}(\bar{\mathbf{v}}))^{-1}\right] \ge \lambda_{\min}^{-1}\left[\tilde{\mathbf{F}}^T(\bar{\mathbf{v}})\tilde{\mathbf{F}}(\bar{\mathbf{v}})\right] > 1/\beta$.

Clearly, β is an important metric for PMU placement from the observability and accuracy perspective. Next, we show that β in fact partially contributes to the metric we propose for PMU placement that also determines the numerical stability and convergence rate.

3.2. Algorithm Convergence

To study the convergence $\lim_{k\to\infty}\mathbf{v}^k = \widehat{\mathbf{v}}$, we prove the following.

Lemma 1. *The error $\left\|\mathbf{v}^{k+1} - \widehat{\mathbf{v}}\right\|$ at the $(k+1)$-th iteration satisfies*

$$\left\|\mathbf{v}^{k+1} - \widehat{\mathbf{v}}\right\| \le \frac{1}{2}\sqrt{\frac{\omega_k}{\beta}}\left\|\mathbf{v}^k - \widehat{\mathbf{v}}\right\|^2 + \frac{\sqrt{2}\epsilon_\star\sqrt{\omega_k}}{\beta}\left\|\mathbf{v}^k - \widehat{\mathbf{v}}\right\|$$

where $\omega_k = \left(\mathbf{v}^k - \widehat{\mathbf{v}}\right)^T\mathbf{M}\left(\mathbf{v}^k - \widehat{\mathbf{v}}\right)/\left\|\mathbf{v}^k - \widehat{\mathbf{v}}\right\|^2$ is a Rayleigh quotient of $\mathbf{M} \triangleq \mathcal{S}(\mathbf{I}_{2N})$ in (10) *and $\epsilon_\star = \|\mathbf{z} - \mathbf{f}(\widehat{\mathbf{v}})\|$.*

Proof. See Appendix B. □

Lemma 1 relates ω_k with the error dynamics. Although ω_k can be easily bounded by the largest eigenvalue of $\mathbf{M}$, this is a pessimistic bound that ignores the dependency of $\mathbf{v}^k$ on $\mathbf{v}^0$ and hence implicitly on the placement $\boldsymbol{\mathcal{V}}$. Without loss of generality, we denote an upper bound ω for all k and discuss how to link ω with $\boldsymbol{\mathcal{V}}$ explicitly later.

Theorem 1. *[29, Theorem 1] Given an upper bound $\omega_k \le \omega$ for all k and suppose $\epsilon_\star\sqrt{2\omega} < \beta$, then $\mathbf{v}^k$ converges $\lim_{k\to\infty}\left\|\mathbf{v}^k - \widehat{\mathbf{v}}\right\| = 0$ as long as $\left\|\mathbf{v}^0(\boldsymbol{\mathcal{V}}) - \widehat{\mathbf{v}}\right\| \le 2\sqrt{\beta/\omega} - 2\sqrt{2}\epsilon_\star/\sqrt{\beta}$.*

Theorem 1 suggests that when ω is large (i.e., the Jacobians exhibit large fluctuations) and β is small (i.e., FIM in (8) is numerically singular), one needs to initialize rather close to $\bar{\mathbf{v}}$ for convergence. Fortunately, we obtain an optimal PMU placement later that can ameliorate the convergence behavior by maximizing the *joint accuracy-convergence (JAC) metric* proposed below

$$\rho_\star = \sqrt{\beta/\omega}, \quad (\text{JAC metric}) \tag{11}$$

where β is defined in (9) and ω is the upper bound in Lemma 1. Note that $1/\rho_\star$ is a measure of the local convergence rate of the algorithm

[1] The derivation is tedious but straightforward from (9) and (23) and thus omitted here due to limited space.

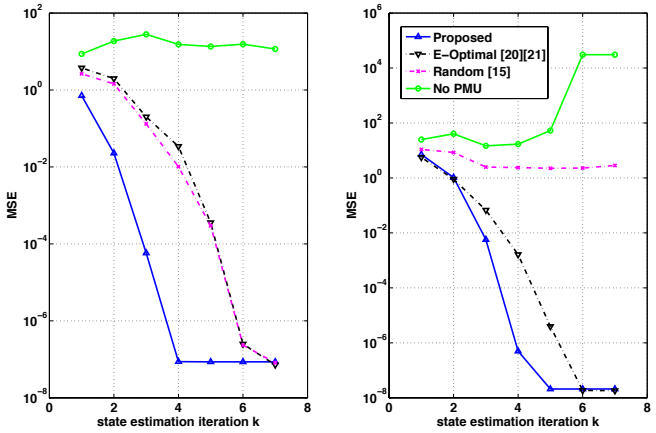

Fig. 1. MSE curves for the IEEE 30 (left) and 118 (right) systems.

when $\epsilon_\star$ is small. Hence, this metric reflects the numerical stability of the algorithm via ω, and the estimation accuracy via β, both of which depend on PMU placement. The greater the *JAC metric*, the more robust to initialization (numerical stability) and the faster the GN algorithm converges. To formulate the PMU placement problem, we examine below how $\boldsymbol{\mathcal{V}}$ affects this *JAC metric* $\rho_\star$ more precisely.

4. OPTIMAL PMU PLACEMENT

4.1. Effects of the PMU Placement on β and ω

We have established the expression of β in relation to $\boldsymbol{\mathcal{V}}$ in (10), which however requires a complicated search over $\mathbf{v} \in \mathbb{V}$. For simplicity, the common practice is to replace the search by substituting a nominal state $\mathbf{v}_{\text{nom}} = [\mathbf{1}^T, \mathbf{0}^T]^T$ as in [21]. On the other hand, the maximum ω requires analyzing the evolution of the Rayleigh quotients ω_k in Lemma 1, which in fact can be easily bounded by the largest eigenvalue of $\mathbf{M}$. However, this is a pessimistic bound that ignores the structure of $\mathbf{v}^k$ resulted from the initializer $\mathbf{v}^0(\boldsymbol{\mathcal{V}})$, which neglects the effects of PMU placement. Assuming that the algorithm makes progress at each iteration such that $\|\mathbf{v}^k - \widehat{\mathbf{v}}\|$ is contracting, a natural choice for ω in the JAC metric is to bound the first Rayleigh quotient ω_0 by an appropriate PMU deployment.

Thus next, we use the upper bound of ω_0 as the surrogate maximum ω for the JAC metric in order to optimize our PMU placement. If the PMUs are accurate, then $\mathbf{z}_{\mathcal{V}} \approx \bar{\mathbf{v}} \approx \widehat{\mathbf{v}}$ and therefore, $\mathbf{v}^0(\boldsymbol{\mathcal{V}}) - \widehat{\mathbf{v}} \approx (\mathbf{I}_{2N} - \mathbf{J}_{\mathcal{V}})(\mathbf{s}_{\mathcal{V}} - \widehat{\mathbf{v}})$, which implies that

$$\omega_0 \approx \frac{\left(\mathbf{v}^0 - \widehat{\mathbf{v}}\right)^T (\mathbf{I}_{2N} - \mathbf{J}_{\mathcal{V}})^T \mathbf{M} (\mathbf{I}_{2N} - \mathbf{J}_{\mathcal{V}}) \left(\mathbf{v}^0 - \widehat{\mathbf{v}}\right)}{(\mathbf{v}^0 - \widehat{\mathbf{v}})^T (\mathbf{I}_{2N} - \mathbf{J}_{\mathcal{V}})^T (\mathbf{I}_{2N} - \mathbf{J}_{\mathcal{V}}) (\mathbf{v}^0 - \widehat{\mathbf{v}})}. \quad (12)$$

By the idempotence $(\mathbf{I}_{2N} - \mathbf{J}_{\mathcal{V}}) = (\mathbf{I}_{2N} - \mathbf{J}_{\mathcal{V}})^2$ in the numerator, the maximum is $\omega \triangleq \lambda_{\max}\left[(\mathbf{I}_{2N} - \mathbf{J}_{\mathcal{V}})^T \mathbf{M} (\mathbf{I}_{2N} - \mathbf{J}_{\mathcal{V}})\right] \geq \omega_0$.

4.2. Problem Formulation and Solution

We have shown the effects of PMU placement $\boldsymbol{\mathcal{V}}$ on ω and β. Thus given a PMU budget N_{PMU}, the *optimal* design aims at maximizing

$$\max_{\boldsymbol{\mathcal{V}}} \quad \frac{\lambda_{\min}\left[\mathcal{P}(\boldsymbol{\mathcal{V}}) + \mathcal{S}(\mathbf{v}_{\text{nom}}\mathbf{v}_{\text{nom}}^T)\right]}{\lambda_{\max}\left[(\mathbf{I}_{2N} - \mathbf{J}_{\mathcal{V}})^T \mathbf{M} (\mathbf{I}_{2N} - \mathbf{J}_{\mathcal{V}})\right]} \quad (13)$$

$$\text{s.t.} \quad \mathbf{J}_{\mathcal{V}} = \mathbf{I}_2 \otimes \text{diag}(\boldsymbol{\mathcal{V}}),\ \mathbf{1}_N^T \boldsymbol{\mathcal{V}} = N_{\text{PMU}},\ \mathcal{V}_n \in \{0,1\}.$$

To avoid solving this complicated eigenvalue problem with integer constraints, we relax (13) by converting the integer constraint $\mathcal{V}_n \in \{0,1\}$ to a convex constraint $0 \leq \mathcal{V}_n \leq 1$, and re-formulate the problem via linear matrix inequalities using two dummy variables

$$\max_{\boldsymbol{\mathcal{V}},\beta,\omega} \beta/\omega,\ \text{s.t.}\ \mathcal{P}(\boldsymbol{\mathcal{V}}) + \mathcal{S}(\mathbf{v}_{\text{nom}}\mathbf{v}_{\text{nom}}^T) \succeq \beta \mathbf{I}_{2N} \quad (14)$$

$$\begin{bmatrix} \omega \mathbf{I}_{2N} & \mathbf{M}^{\frac{1}{2}}(\mathbf{I}_{2N} - \mathbf{J}_{\mathcal{V}}) \\ (\mathbf{I}_{2N} - \mathbf{J}_{\mathcal{V}})^T \mathbf{M}^{\frac{T}{2}} & \mathbf{I}_{2N} \end{bmatrix} \succeq \mathbf{0}$$

$$\mathbf{J}_{\mathcal{V}} = \mathbf{I}_2 \otimes \text{diag}(\boldsymbol{\mathcal{V}}),\ \mathbf{1}_N^T \boldsymbol{\mathcal{V}} = N_{\text{PMU}}, \boldsymbol{\mathcal{V}} \in [0,1]^{2N}$$

In general, this is a quasi-convex problem that can be solved in a globally optimal fashion via the classical bisection method by performing a sequence of semidefinite programs (SDP) feasibility problems [30]. Fortunately, since the objective (14) is a linear fractional function, one can use the *Charnes-Cooper* transformation [31], and re-formulate (14) equivalently as a single SDP instead.

Proposition 1. *By letting $\gamma = 1/\omega$, $\kappa = \beta/\omega$ and $\boldsymbol{\xi} = \boldsymbol{\mathcal{V}}/\omega$, the global optimum solution to* (14) *can be determined by*

$$\max_{\boldsymbol{\xi},\kappa,\gamma} \kappa \quad \text{s.t.}\ \mathcal{P}(\boldsymbol{\xi}) + \gamma \mathcal{S}(\mathbf{v}_{nom}\mathbf{v}_{nom}^T) \succeq \kappa \mathbf{I}_{2N} \quad (15)$$

$$\begin{bmatrix} \mathbf{I}_{2N} & \mathbf{M}^{\frac{1}{2}}(\gamma \mathbf{I}_{2N} - \mathbf{I}_{\boldsymbol{\xi}}) \\ (\gamma \mathbf{I}_{2N} - \mathbf{I}_{\boldsymbol{\xi}})^T \mathbf{M}^{\frac{T}{2}} & \gamma \mathbf{I}_{2N} \end{bmatrix} \succeq \mathbf{0}$$

$$\mathbf{I}_{\boldsymbol{\xi}} = \mathbf{I}_2 \otimes \text{diag}(\boldsymbol{\xi}),\ \mathbf{1}_N^T \boldsymbol{\xi} = N_{PMU}\gamma,\ \mathbf{0}_N \preceq \boldsymbol{\xi} \preceq \gamma \mathbf{1}_N,$$

whose solution is mapped to the problem in (14) *via $\boldsymbol{\mathcal{V}}_\star = \boldsymbol{\xi}_\star/\gamma_\star$.*

The solution above may not have binary entries. Therefore we set the largest N_{PMU} values in $\boldsymbol{\mathcal{V}}_\star$ to 1 and others to 0 as in [21]. Note that this optimization is solved once off-line, and the effects of such placement on the convergence and accuracy of the state estimation is illustrated in the simulations presented next.

5. SIMULATIONS

We illustrate the convergence and MSE performance using different placements, where $\text{MSE} = \mathbb{E}\|\mathbf{v}^k - \widehat{\mathbf{v}}\|^2/N$ for each iteration k. We compare our *optimal* design against a *random* placement and the *E-optimal* scheme[2] in [20, 21] that solely optimizes the estimation accuracy via the FIM in (8). The simulation uses the IEEE 30 and 118 systems in MATPOWER 4.0. The measurements are generated with independent errors $\mathbf{R} = \sigma^2 \mathbf{I}$ and $\sigma^2 = 10^{-4}$ and we use 50% of all SCADA measurements chosen at random.

We compare the MSE curves against the state estimation iteration k for each placement with $N_{\text{PMU}} = 5$ for the IEEE-30 bus system and $N_{\text{PMU}} = 20$ for the IEEE-118 system (17% installation) in Fig. 1. To verify the robustness to numerical stability and the convergence rate, the MSE curves are averaged over 100 runs. For each run, we generate an independent placement for the *random* scheme that guarantees system observability [15], and use a non-informative initializer $\mathbf{s}_{\mathcal{V}} = [\mathbf{1}_N^T + 0.1\boldsymbol{\varepsilon}^T, \mathbf{0}_N^T]^T$ perturbed by a Gaussian random vector $\boldsymbol{\varepsilon}$ with a covariance $\mathbb{E}\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\right] = \mathbf{I}_{2N}$. We keep the imaginary part unperturbed considering that voltage phases are usually small.

It is seen in Fig. 1 that if there are no PMU installed, it is possible that the algorithm does not converge due to bad initializations. In contrast, the proposed scheme converges stably to the ML estimates even under 10% perturbation. The performance of *random* placement is not stably guaranteed because it diverges for the large scale 118-bus system in Fig. 1. Consistent with Theorem 1, since the

[2]Among the *A*, *M*, *D-optimal* designs in [20, 21], we choose *E-optimal* design because of the common objective in maximizing β. These designs provide similar performances to *E-optimal* and hence are not repeated.

noise variance σ^2 is small, the algorithm asymptotically converges quadratically for the *optimal* and the *E-optimal* placement, but the convergence rates are very different in both figures. Although the MSE performances after convergence remain comparable, the *optimal* placement considerably accelerates the convergence compared to the *random* and *E-optimal* placement.

6. CONCLUSIONS

In this paper, we propose a useful metric, referred to as JAC, to evaluate the convergence and accuracy of hybrid PSSE for a given sensor deployment, where PMUs are used to initialize the Gauss-Newton iterative estimation. We optimize our placement strategy with respect to the JAC metric via a simple SDP, and confirm numerically the convergence and estimation performance of the proposed scheme.

A. POWER FLOW EQUATIONS AND JACOBIAN MATRIX

The matrix $\mathbf{Y} = [-Y_{nm}]$ includes line admittances $Y_{nm} = G_{nm} + \mathrm{i}B_{nm}$, $(n,m) \in \mathcal{E}$ and shunt admittances $\bar{Y}_{nm} = \bar{G}_{nm} + \mathrm{i}\bar{B}_{nm}$ of the line $(n,m) \in \mathcal{E}$, and self-admittance $Y_{nn} = -\sum_{m \neq n}(\bar{Y}_{nm} + Y_{nm})$. Using $\mathbf{e}_n = [0, \cdots, 1, \cdots, 0]^T$, we define the following

$$\mathbf{Y}_n \triangleq \mathbf{e}_n\mathbf{e}_n^T\mathbf{Y},\ \mathbf{Y}_{nm} \triangleq (Y_{nm} + \bar{Y}_{nm})\mathbf{e}_n\mathbf{e}_n^T - Y_{nm}\mathbf{e}_n\mathbf{e}_l^T. \quad (16)$$

Letting $\mathbf{G}_n = \Re\{\mathbf{Y}_n\}$, $\mathbf{B}_n = \Im\{\mathbf{Y}_n\}$, $\mathbf{G}_{nm} = \Re\{\mathbf{Y}_{nm}\}$ and $\mathbf{B}_{nm} = \Im\{\mathbf{Y}_{nm}\}$, we further define the following matrices

$$\mathbf{N}_{P,n} \triangleq \begin{bmatrix} \mathbf{G}_n & -\mathbf{B}_n \\ \mathbf{B}_n & \mathbf{G}_n \end{bmatrix} \qquad \mathbf{N}_{Q,n} \triangleq -\begin{bmatrix} \mathbf{B}_n & \mathbf{G}_n \\ -\mathbf{G}_n & \mathbf{B}_n \end{bmatrix}$$

$$\mathbf{E}_{P,nm} \triangleq \begin{bmatrix} \mathbf{G}_{nm} & -\mathbf{B}_{nm} \\ \mathbf{B}_{nm} & \mathbf{G}_{nm} \end{bmatrix} \ \mathbf{E}_{Q,nm} \triangleq -\begin{bmatrix} \mathbf{B}_{nm} & \mathbf{G}_{nm} \\ -\mathbf{G}_{nm} & \mathbf{B}_{nm} \end{bmatrix}$$

$$\mathbf{C}_{I,nm} \triangleq \begin{bmatrix} \mathbf{G}_{nm} & \mathbf{0} \\ \mathbf{0} & -\mathbf{B}_{nm} \end{bmatrix} \ \mathbf{C}_{J,nm} \triangleq \begin{bmatrix} \mathbf{B}_{nm} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_{nm} \end{bmatrix}.$$

The SCADA systems collect the active/reactive power injection (P_n, Q_n) at bus n and flow (P_{nm}, Q_{nm}) at bus n from line (n,m)

$$P_n = \mathbf{v}^T\mathbf{N}_{P,n}\mathbf{v}, \quad Q_n = \mathbf{v}^T\mathbf{N}_{Q,n}\mathbf{v}, \quad (17)$$

$$P_{nm} = \mathbf{v}^T\mathbf{E}_{P,nl}\mathbf{v}, \quad Q_{nm} = \mathbf{v}^T\mathbf{E}_{Q,nl}\mathbf{v}. \quad (18)$$

We stack these functions in the power flow equation vectors

$$\mathbf{f}_{\mathcal{I}}(\mathbf{v}) = [\cdots, P_n, \cdots, \cdots, Q_n, \cdots]^T \quad (19)$$

$$\mathbf{f}_{\mathcal{F}}(\mathbf{v}) = [\cdots, P_{nm}, \cdots, \cdots, Q_{nm}, \cdots]^T. \quad (20)$$

The WAMS collects the voltage phasor $(\Re\{V_n\}, \Im\{V_n\})$ at bus n and the current phasor (I_{nm}, J_{nm}) at bus n on line (n,m)

$$I_{nm} = (\mathbf{1}_2 \otimes \mathbf{e}_n)^T\mathbf{C}_{I,nl}\mathbf{v},\ J_{nm} = (\mathbf{1}_2 \otimes \mathbf{e}_n)^T\mathbf{C}_{J,nl}\mathbf{v}, \quad (21)$$

where $\otimes$ is the Kronecker product, and stacks them as

$$\mathbf{f}_{\mathcal{V}}(\mathbf{v}) = \mathbf{v}, \quad \mathbf{f}_{\mathcal{C}}(\mathbf{v}) = [\cdots, I_{nm}, \cdots, \cdots, J_{nm}, \cdots]^T. \quad (22)$$

The Jacobian $\mathbf{F}(\mathbf{v})$ can then be derived from (21), (17) and (18) as

$$\mathbf{F}(\mathbf{v}) = \begin{bmatrix} \mathbf{I}_{2N} & \mathbf{H}_{\mathcal{C}}^T & \mathbf{H}_{\mathcal{I}}^T(\mathbf{I}_{2N} \otimes \mathbf{v}) & \mathbf{H}_{\mathcal{F}}^T(\mathbf{I}_{4L} \otimes \mathbf{v}) \end{bmatrix}^T, \quad (23)$$

where $\mathbf{H}_{\mathcal{I}} \triangleq \left[\cdots, \mathbf{N}_{P,n} + \mathbf{N}_{P,n}^T, \cdots, \mathbf{N}_{Q,n} + \mathbf{N}_{Q,n}^T, \cdots\right]^T$, and $\mathbf{H}_{\mathcal{F}} \triangleq \left[\cdots, \mathbf{E}_{P,nl} + \mathbf{E}_{P,nl}^T, \cdots, \mathbf{E}_{Q,nl} + \mathbf{E}_{Q,nl}^T, \cdots\right]^T$ and $\mathbf{H}_{\mathcal{C}} \triangleq \left[\cdots, \mathbf{H}_{I,n}^T, \cdots, \cdots, \mathbf{H}_{J,n}^T, \cdots\right]^T$, $\mathbf{H}_{I,n} \triangleq \mathbf{S}_n\mathbf{C}_{I,n}$ with $\mathbf{C}_{I,n} \triangleq \left[\cdots, \mathbf{C}_{I,nl}^T, \cdots\right]^T$ and $\mathbf{H}_{J,n} \triangleq \mathbf{S}_n\mathbf{C}_{J,n}$ with $\mathbf{C}_{J,n} \triangleq \left[\cdots, \mathbf{C}_{J,nl}^T, \cdots\right]^T$ using $\mathbf{S}_n \triangleq \mathbf{I}_{L_n} \otimes (\mathbf{1}_2 \otimes \mathbf{e}_n)^T$.

B. PROOF OF LEMMA 1

We first prove that the Jacobian $\tilde{\mathbf{F}}(\mathbf{v})$ satisfies $\left\|\tilde{\mathbf{F}}(\mathbf{v}) - \tilde{\mathbf{F}}(\mathbf{v}')\right\|^2 \leq (\mathbf{v} - \mathbf{v}')^T\mathbf{M}(\mathbf{v} - \mathbf{v}')$ with $\mathbf{M} = \mathcal{S}(\mathbf{I}_{2N})$ defined in (10). Using

$$\tilde{\mathbf{F}}(\mathbf{v}) - \tilde{\mathbf{F}}(\mathbf{v}') = \begin{bmatrix} \mathbf{0}_{2N\times 2N} \\ \mathbf{0}_{4L\times 2N} \\ [\mathbf{R}_{\mathcal{I}}^{-\frac{1}{2}}\mathbf{J}_{\mathcal{I}} \otimes (\mathbf{v} - \mathbf{v}')^T]\mathbf{H}_{\mathcal{I}} \\ [\mathbf{R}_{\mathcal{F}}^{-\frac{1}{2}}\mathbf{J}_{\mathcal{F}} \otimes (\mathbf{v} - \mathbf{v}')^T]\mathbf{H}_{\mathcal{F}} \end{bmatrix} \quad (24)$$

together with the norm inequality $\|\cdot\| \leq \|\cdot\|_F$ and $\|\cdot\|_F^2 = \mathrm{Tr}\left((\cdot)^T(\cdot)\right)$ and the property of Kronecker products, we have

$$\begin{aligned}\left\|\tilde{\mathbf{F}}(\mathbf{v}) - \tilde{\mathbf{F}}(\mathbf{v}')\right\|_F^2 &= \mathrm{Tr}\left[\mathbf{H}_{\mathcal{I}}^T\left(\mathbf{J}_{\mathcal{I}}\mathbf{R}_{\mathcal{I}}^{-1}\mathbf{J}_{\mathcal{I}}^T \otimes (\mathbf{v} - \mathbf{v}')(\mathbf{v} - \mathbf{v}')^T\right)\mathbf{H}_{\mathcal{I}}\right] \\ &\quad + \mathrm{Tr}\left[\mathbf{H}_{\mathcal{F}}^T\left(\mathbf{J}_{\mathcal{F}}\mathbf{R}_{\mathcal{F}}^{-1}\mathbf{J}_{\mathcal{F}}^T \otimes (\mathbf{v} - \mathbf{v}')(\mathbf{v} - \mathbf{v}')^T\right)\mathbf{H}_{\mathcal{F}}\right].\end{aligned}$$

Using the property of trace operators, the sub-matrices of $\mathbf{H}_{\mathcal{I}}$ and $\mathbf{H}_{\mathcal{F}}$ in (23), we can express the above norm by expanding the Kronecker product $\otimes$ and re-arrange the summation into $\mathbf{M} = \mathcal{S}(\mathbf{I}_{2N})$

$$\left\|\tilde{\mathbf{F}}(\mathbf{v}) - \tilde{\mathbf{F}}(\mathbf{v}')\right\|_F^2 = \mathrm{Tr}\left[(\mathbf{v} - \mathbf{v}')(\mathbf{v} - \mathbf{v}')^T\mathbf{M}\right] \quad (25)$$

$$= (\mathbf{v} - \mathbf{v}')^T\mathbf{M}(\mathbf{v} - \mathbf{v}'). \quad (26)$$

Using this result we are now ready to prove Lemma 1. We follow the same derivations in [29, Lem. 2] to obtain

$$\begin{aligned}\mathbf{v}^{k+1} - \widehat{\mathbf{v}} &= \tilde{\mathbf{F}}^\dagger(\mathbf{v}^k)\left[\tilde{\mathbf{F}}(\mathbf{v}^k)\left(\mathbf{v}^k - \widehat{\mathbf{v}}\right) + \tilde{\mathbf{f}}(\widehat{\mathbf{v}}) - \tilde{\mathbf{f}}(\mathbf{v}^k)\right] \\ &\quad + \left[\tilde{\mathbf{F}}^\dagger(\mathbf{v}^k) - \tilde{\mathbf{F}}^\dagger(\widehat{\mathbf{v}})\right]\left(\tilde{\mathbf{c}} - \tilde{\mathbf{f}}(\widehat{\mathbf{v}})\right). \quad (27)\end{aligned}$$

The first term can be written with the mean-value theorem and (23)

$$\begin{aligned}&\tilde{\mathbf{F}}(\mathbf{v}^k)\left(\mathbf{v}^k - \widehat{\mathbf{v}}\right) + \tilde{\mathbf{f}}(\widehat{\mathbf{v}}) - \tilde{\mathbf{f}}(\mathbf{v}^k) \\ &= \int_0^1 \left[\tilde{\mathbf{F}}(\mathbf{v}^k) - \tilde{\mathbf{F}}\left(\widehat{\mathbf{v}} + t(\mathbf{v}^k - \widehat{\mathbf{v}})\right)\right]\left(\mathbf{v}^k - \widehat{\mathbf{v}}\right)\mathrm{d}t \\ &= \frac{1}{2}\left[\tilde{\mathbf{F}}(\mathbf{v}^k) - \tilde{\mathbf{F}}(\widehat{\mathbf{v}})\right]\left(\mathbf{v}^k - \widehat{\mathbf{v}}\right).\end{aligned}$$

Then the first term in (27) can be bounded as

$$\|\tilde{\mathbf{F}}(\mathbf{v}^k)\left(\mathbf{v}^k - \widehat{\mathbf{v}}\right) + \tilde{\mathbf{f}}(\widehat{\mathbf{v}}) - \tilde{\mathbf{f}}(\mathbf{v}^k)\| \leq \frac{1}{2}\|\tilde{\mathbf{F}}(\mathbf{v}^k) - \tilde{\mathbf{F}}(\widehat{\mathbf{v}})\|\|\mathbf{v}^k - \widehat{\mathbf{v}}\|.$$

From [32, Lem. 1], part of the second term in (27) is bounded as

$$\|\tilde{\mathbf{F}}^\dagger(\mathbf{v}^k) - \tilde{\mathbf{F}}^\dagger(\widehat{\mathbf{v}})\| \leq \sqrt{2}\|\tilde{\mathbf{F}}^\dagger(\mathbf{v}^k)\|\|\tilde{\mathbf{F}}^\dagger(\widehat{\mathbf{v}})\|\|\tilde{\mathbf{F}}(\mathbf{v}^k) - \tilde{\mathbf{F}}(\widehat{\mathbf{v}})\|.$$

Using the result in (25), by substituting the above bounds back to (27) and letting $\epsilon_\star = \|\mathbf{z} - \mathbf{f}(\widehat{\mathbf{v}})\|$, the recursion (27) is bounded as

$$\left\|\mathbf{v}^{k+1} - \widehat{\mathbf{v}}\right\| \leq \left(\frac{\left\|\mathbf{v}^k - \widehat{\mathbf{v}}\right\|}{2\sqrt{\beta}} + \frac{\sqrt{2}\epsilon_\star}{\beta}\right)\sqrt{(\mathbf{v}^k - \widehat{\mathbf{v}})^T\mathbf{M}(\mathbf{v}^k - \widehat{\mathbf{v}})}.$$

Let $\omega_k = \left(\mathbf{v}^k - \widehat{\mathbf{v}}\right)^T\mathbf{M}\left(\mathbf{v}^k - \widehat{\mathbf{v}}\right)/\left\|\mathbf{v}^k - \widehat{\mathbf{v}}\right\|^2$ be the Rayleigh quotient of $\mathbf{M}$ in iteration k. Then the quadratic form becomes $(\mathbf{v}^k - \widehat{\mathbf{v}})^T\mathbf{M}(\mathbf{v}^k - \widehat{\mathbf{v}}) = \omega_k\left\|\mathbf{v}^k - \widehat{\mathbf{v}}\right\|^2$, which leads to Lemma 1.

C. REFERENCES

[1] J. Chow, P. Quinn, L. Beard, D. Sobajic, A. Silverstein, and L. Vanfretti, "Guidelines for Siting Phasor Measurement Units: Version 8, June 15, 2011 North American SynchroPhasor Initiative (NASPI) Research Initiative Task Team (RITT) Report," 2011.

[2] T. Yang, H. Sun, and A. Bose, "Transition to a Two-Level Linear State Estimator : Part I & II," *IEEE Trans. Power Syst.*, no. 99, pp. 1–1, 2011.

[3] N. Manousakis, G. Korres, and P. Georgilakis, "Optimal Placement of Phasor Measurement Units: A Literature Review," in *Intelligent System Application to Power Systems (ISAP), 2011 16th International Conference on*. IEEE, 2011, pp. 1–6.

[4] W. Yuill, A. Edwards, S. Chowdhury, and S. Chowdhury, "Optimal PMU Placement: A Comprehensive Literature Review," in *Power and Energy Society General Meeting, 2011 IEEE*. IEEE, 2011, pp. 1–8.

[5] B. Singh, N. Sharma, A. Tiwari, K. Verma, and S. Singh, "Applications of Phasor Measurement Units (PMUs) in Electric Power System Networks Incorporated with FACTS Controllers," *International Journal of Engineering, Science and Technology*, vol. 3, no. 3, 2011.

[6] K. Clements, "Observability Methods and Optimal Meter Placement," *International Journal of Electrical Power & Energy Systems*, vol. 12, no. 2, pp. 88–93, 1990.

[7] X. Bei, Y. Yoon, and A. Abur, "Optimal Placement and Utilization of Phasor Measurements for State Estimation," *PSERC Publication*, pp. 05–20, 2005.

[8] R. Nuqui and A. Phadke, "Phasor Measurement Unit Placement Techniques for Complete and Incomplete Observability," *Power Delivery, IEEE Trans.*, vol. 20, no. 4, pp. 2381–2388, 2005.

[9] D. Dua, S. Dambhare, R. Gajbhiye, and S. Soman, "Optimal Multistage Scheduling of PMU Placement: An ILP Approach," *Power Delivery, IEEE Trans.*, vol. 23, no. 4, pp. 1812–1820, 2008.

[10] R. Emami and A. Abur, "Robust Measurement Design by Placing Synchronized Phasor Measurements on Network Branches," *Power Systems, IEEE Trans.*, vol. 25, no. 1, pp. 38–43, 2010.

[11] B. Gou, "Optimal Placement of PMUs by Integer Linear Programming," *Power Systems, IEEE Trans.*, vol. 23, no. 3, pp. 1525–1526, 2008.

[12] N. Abbasy and H. Ismail, "A Unified Approach for the Optimal PMU Location for Power System State Estimation," *Power Systems, IEEE Trans.*, vol. 24, no. 2, pp. 806–813, 2009.

[13] S. Chakrabarti and E. Kyriakides, "Optimal Placement of Phasor Measurement Units for Power System Observability," *Power Systems, IEEE Trans.*, vol. 23, no. 3, pp. 1433–1440, 2008.

[14] S. Chakrabarti, E. Kyriakides, and D. Eliades, "Placement of Synchronized Measurements for Power System Observability," *Power Delivery, IEEE Trans.*, vol. 24, no. 1, pp. 12–19, 2009.

[15] T. Baldwin, L. Mili, M. Boisen Jr, and R. Adapa, "Power System Observability with Minimal Phasor Measurement Placement," *Power Systems, IEEE Trans.*, vol. 8, no. 2, pp. 707–715, 1993.

[16] B. Milosevic and M. Begovic, "Nondominated Sorting Genetic Algorithm for Optimal Phasor Measurement Placement," *Power Systems, IEEE Trans.*, vol. 18, no. 1, pp. 69–75, 2003.

[17] F. Aminifar, C. Lucas, A. Khodaei, and M. Fotuhi-Firuzabad, "Optimal Placement of Phasor Measurement Units Using Immunity Genetic Algorithm," *Power Delivery, IEEE Trans.*, vol. 24, no. 3, pp. 1014–1020, 2009.

[18] S. Chakrabarti and E. Kyriakides, "PMU Measurement Uncertainty Considerations in WLS State Estimation," *Power Systems, IEEE Trans.*, vol. 24, no. 2, pp. 1062–1071, 2009.

[19] J. Zhang, G. Welch, and G. Bishop, "Observability and Estimation Uncertainty Analysis for PMU Placement Alternatives," in *North American Power Symposium (NAPS), 2010*. IEEE, 2010, pp. 1–8.

[20] Q. Li, R. Negi, and M. Ilic, "Phasor Measurement Units Placement for Power System State Estimation: a Greedy Approach," in *Power and Energy Society General Meeting, 2011 IEEE*. IEEE, 2011, pp. 1–8.

[21] V. Kekatos and G. Giannakis, "A Convex Relaxation Approach to Optimal Placement of Phasor Measurement Units," in *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2011 4th IEEE International Workshop on*. IEEE, 2011, pp. 145–148.

[22] K. Zhu, L. Nordstrom, and L. Ekstam, "Application and Analysis of Optimum PMU Placement Methods with Application to State Estimation Accuracy," in *Power & Energy Society General Meeting, 2009. PES'09. IEEE*. Ieee, 2009, pp. 1–7.

[23] M. Asprou and E. Kyriakides, "Optimal PMU Placement for Improving Hybrid State Estimator Accuracy," in *PowerTech, 2011 IEEE Trondheim*. IEEE, 2011, pp. 1–7.

[24] Q. Li, T. Cui, Y. Weng, R. Negi, F. Franchetti, and M. Ilic, "An Information-Theoretic Approach to PMU Placement in Electric Power Systems," *Arxiv preprint arXiv:1201.2934*, 2012.

[25] S. Pajic and K. Clements, "Power System State Estimation via Globally Convergent Methods," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1683–1689, 2005.

[26] F. Schweppe and E. Handschin, "Static State Estimation in Electric Power Systems," *Proceedings of the IEEE*, vol. 62, no. 7, pp. 972–982, 1974.

[27] R. Larson, W. Tinney, and J. Peschon, "State Estimation in Power Systems Part I: Theory and Feasibility," *IEEE Trans. Power App. Syst.*, no. 3Part-I, pp. 345–352, 1970.

[28] B. Stott and O. Alsaç, "Fast Decoupled Load Flow," *IEEE Trans. Power App. Syst.*, no. 3, pp. 859–869, 1974.

[29] X. Li and A. Scaglione, "Convergence and Applications of a Gossip-based Gauss-Newton Algorithm," *submitted to IEEE Trans. on Signal Processing, preprint arXiv:1210.0056*, 2012.

[30] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge Univ Pr, 2004.

[31] A. Charnes and W. Cooper, "Programming with Linear Fractional Functionals," *Naval Research logistics quarterly*, vol. 9, no. 3-4, pp. 181–186, 1962.

[32] S. Salzo and S. Villa, "Convergence Analysis of a Proximal Gauss-Newton Method," *Arxiv preprint arXiv:1103.0414*, 2011.