Optimal Placement of Phasor Measurement Units via Convex Relaxation

Vassilis Kekatos, *Member, IEEE*, Georgios B. Giannakis, *Fellow, IEEE*, and Bruce Wollenberg, *Life Fellow, IEEE*

***Abstract*—Instrumenting power networks with phasor measurement units (PMUs) facilitates several tasks including optimum power flow, system control, contingency analysis, visualization, and integration of renewable resources, thus enabling situational awareness—one of the key steps toward realizing the smart grid vision. The installation cost of PMUs currently prohibits their deployment on every bus, which in turn motivates their strategic placement across the power grid. As state estimation is at the core of grid monitoring, PMU deployment is optimized here based on estimation-theoretic criteria. Considering both voltage and incident current readings per PMU-instrumented bus and incorporating conventionally derived state estimates under the Bayesian framework, PMU placement is formulated as an optimal experimental design task. To bypass the combinatorial search involved, a convex relaxation is developed to obtain solutions with numerical optimality guarantees. In the tests performed on standard IEEE 14-, 30-, and 118-bus benchmarks, the proposed relaxation approaches and oftentimes attains the optimum PMU placement.**

***Index Terms*—Gradient projection method, maximum a-posteriori estimation, optimal experimental design, phasor measurement units, SCADA measurements, semidefinite programming.**

I. Introduction

PHASOR measurement units (PMUs) are contemporary metering devices installed on system buses to measure phasors of bus voltages and currents flowing across lines [14], [31]. After sampling, windowing, phasor estimation, and time stamping, measurements are communicated to the intended application through phasor data concentrators. Merits of PMUs (a.k.a. synchrophasors) over conventional power meters include increased precision in measuring phasor angles due to network-wide synchronization, and higher sampling rates.

PMU penetration has so far been rather limited, mainly due to the installation and networking costs involved [34]. However, their important role in network operation and the growing maturity of PMU technology are expected to markedly increase their deployment [9]. According to the North American Synchrophasor Initiative [13], the number of PMUs installed and networked in the eastern/western interconnection is expected to raise from 105/56 as of 2009 to 400–600 by 2014 [29]. Thus, strategic placement of synchrophasors is currently a critical issue for the power operators worldwide.

Given the gamut of PMU-enabled applications [31], [34], selecting the buses to be PMU-instrumented faces many options to choose from. A multitude of diverse attributes and target applications (along with associated criteria) include observability, oscillation and angular separation monitoring, infrastructure criticality and cyber-security under communication and deployment cost constraints [25], and offer representative options to consider when deciding PMU placement. In the present paper, PMU sites are selected to minimize the error in estimating the grid state. The emphasis is on state estimation, since this task affects directly several critical power monitoring applications, e.g., optimal power flow, contingency analysis, and topology inference. In addition, the proposed methodology can be combined with other placement criteria in a multi-objective optimization framework.

Given values of power, voltages, and currents metered at specific buses and lines of the grid, state estimation amounts to finding the underlying complex voltages at all buses. Traditionally, state estimation has relied on the supervisory control and data acquisition (SCADA) meter readings [28], [37, Ch. 12]. State estimation using PMU measurements was introduced in [32], and the improved estimation accuracy when PMU and SCADA measurements are used jointly has been documented in [39]. Note that with a relatively small number of synchrophasor units, state estimation based solely on PMU readings does not guarantee identifiability; in any case, the SCADA-based information should not be discarded.

Similar to [23], PMU placement is formulated here as a variation of the optimal experimental design problem [33], [7, Sec. 7.5]. Recovering the system state expressed in rectangular coordinates using PMU measurements leads to a linear estimation problem, with the SCADA-based estimate serving the role of a Gaussian prior (Section II). The optimal PMU placement is the one achieving the smallest covariance matrix for the associated state estimation error (Section III). The combinatorial problem involved is suboptimally solved after relaxing it to a convex semidefinite program (SDP) [7], [21]. The SDP-derived cost values provide numerical bounds on the suboptimality gap. After expressing the input data of the problem in terms of the physical properties of the grid in Section IV, numerical tests are provided in Section V. Numerical tests are conducted using standard SDP solvers as well as a customized gradient projection algorithm (cf. Appendix). Surprisingly, the results obtained

Manuscript received August 04, 2011; revised December 07, 2011 and January 20, 2012; accepted January 23, 2012. Date of publication February 24, 2012; date of current version July 18, 2012. The work of V. Kekatos was supported by a Marie Curie International Outgoing Fellowship within the 7-th European Community Framework Programme (No. 234914). This work was supported by NSF grants CCF-1016605, and ECCS-1002180. Paper no. TPWRS-00730-2011.

V. Kekatos is with the Electrical and Computer Engineering Department, University of Minnesota, Minneapolis, MN 55455 USA, and also with the Computer Engineering and Informatics Department, University of Patras, Patras, Greece (e-mail: kekatos@umn.edu).

G. B. Giannakis and B. Wollenberg are with the Electrical and Computer Engineering Department, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: georgios@umn.edu; wollenbe@umn.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRS.2012.2185959

demonstrate that the gap is small and oftentimes zero for the IEEE 14-, 30-, and 118-bus power network benchmarks [36].

Related Work: A large volume of the existing literature on PMU placement strategies targets topological observability of the power network [5]. The latter ensures existence of a spanning tree covering all nodes (buses) with edges (transmission lines) whose currents can be (in)directly metered by PMUs. Using topological observability as a criterion, optimal PMU positions have been pursued via simulated annealing [5], or genetic algorithms [3], [27]. If the number of available PMUs cannot provide a spanning tree, only incomplete observability can be afforded, and a tree search yields the PMU positions sought [30]. By extending topological (in)observability to account for branch-only units, contingencies, and other limitations, the PMU placement problem has been posed as a binary linear program [1], [15], [16], [19], or, as a binary quadratic one [10], [11]; see also [2] for a related probabilistic approach.

In theory, topological observability does not imply observability [5], that is recoverability of the underlying state vector using noise-free linear observations. In practical scenarios though, the implication does hold [12]. But observability does not guarantee meaningful estimation accuracy: a full-rank, yet ill-conditioned linear regression matrix can yield numerically unstable estimators and amplify the noise when dealing with realistic noisy measurements. Hence, estimation accuracy, which considers observability together with the noise statistics, is a more meaningful requirement for the power network.

A hybrid approach targeting topological observability at the first stage, and high estimation accuracy at the second stage in an ad hoc manner, is presented in [38]. In [4], [18], [40], and [23], PMUs are deployed to provide high estimation accuracy. However, the state vector is expressed in polar coordinates, and hence, the informative line current PMU measurements are ignored. Specifically, [18] proposes three placement criteria, one of which is similar to the D-optimal design criterion (cf. Section III), and is subsequently solved by a genetic algorithm with no optimality guarantees. In [4], the state covariance matrix is approximated by Monte Carlo simulations first; then, the bus whose sum of variances over itself and its neighboring buses is maximum is heuristically added to the placement; and the two steps are repeated. Similarly, an incremental algorithm selects the PMU with the highest variance at each step in [40]. Finally, different from the novel approach here, the method in [23] ignores the PMU line current measurements, and estimates only the phase of the system state using a greedy (incremental) approach.

Notation: Lower- (upper-) case boldface letters denote column vectors (matrices), and calligraphic letters stand for sets; $\mathbf{1}_N$ ($\mathbf{0}_N$) is the all-ones(zeros) vector of length N; $(\cdot)^T$ and $(\cdot)^\star$ denote transposition and complex conjugation, respectively; $\|\mathbf{x}\|^2_{\mathbf{A}}$ stands for $\mathbf{x}^T\mathbf{A}\mathbf{x}$; $\mathbf{A} \succeq \mathbf{0}$ designates a semi-positive definite matrix, i.e., a symmetric matrix with non-negative eigenvalues; $\mathcal{N}(\mathbf{m}, \mathbf{\Sigma})$ stands for the multivariate Gaussian probability density function with mean $\mathbf{m}$ and covariance matrix $\mathbf{\Sigma}$.

II. PMU-Based State Estimation

Consider a power grid comprising N_b buses (nodes) connected through N_l transmission lines (edges). After expressing the complex voltages at all buses in rectangular coordinates, the grid state is described by the $2N_b \times 1$ vector $\mathbf{v}_o := [\mathbf{v}_{o,r}^T\ \mathbf{v}_{o,i}^T]^T$, where $\mathbf{v}_{o,r}$ and $\mathbf{v}_{o,i}$ denote the real and imaginary parts of nodal voltages, respectively.

In power transmission systems, each bus is connected to a limited number of lines. Contemporary phasor measurement devices feature an adequate number of channels. Thus, it is reasonable to consider that once the nth bus is selected for PMU instrumentation, the installed unit records not only the complex voltage of bus n, but also the complex currents flowing over all the L_n lines incident to this bus. The synchrophasor measurements at the nth bus expressed in rectangular coordinates are collected in a vector $\mathbf{z}_n \in \mathbb{R}^{M_n}$ with $M_n := 2(L_n + 1)$, and obey the following linear model:

$$\mathbf{z}_n = \mathbf{H}_n\mathbf{v}_o + \mathbf{w}_n \tag{1}$$

where $\mathbf{H}_n \in \mathbb{R}^{M_n \times 2N_b}$ is the associated regression matrix, and $\mathbf{w}_n \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_n)$ denotes the additive Gaussian noise vector that is assumed independent across PMUs. The detailed form of $\mathbf{H}_n$'s will be specified in Section IV-A.

To capture presence or absence of a PMU, a binary variable a_n is introduced per bus: its value is 1 if a PMU is present at the nth bus; and 0, otherwise. For the linear-Gaussian model of (1) and a given PMU indicator vector $\mathbf{a} := [a_1\ \cdots\ a_{N_b}]^T$, the state of the power grid can be deduced using the PMU measurements. Given that the likelihood of the measurements is $p\left(\{\mathbf{z}_n\}_{n=1}^{N_b}; \mathbf{v}_o\right) = \prod_{n=1}^{N_b} p(\mathbf{z}_n; \mathbf{v}_o)^{a_n}$, the maximum likelihood estimate (MLE) of the system state is simply

$$\hat{\mathbf{v}}_{\text{MLE}} := \arg\min_{\mathbf{v}}\ \frac{1}{2}\sum_{n=1}^{N_b} a_n\|\mathbf{z}_n - \mathbf{H}_n\mathbf{v}\|^2_{\mathbf{\Sigma}_n^{-1}}. \tag{2}$$

When the so-called *gain matrix* [28]

$$\mathbf{G} := \sum_{n=1}^{N_b} a_n\mathbf{H}_n^T\mathbf{\Sigma}_n^{-1}\mathbf{H}_n \tag{3}$$

is non-singular, the minimizer of (2) is unique and it is given in closed form as $\hat{\mathbf{v}}_{\text{MLE}} = \mathbf{G}^{-1}\left(\sum_{n=1}^{N_b} a_n\mathbf{H}_n^T\mathbf{\Sigma}_n^{-1}\mathbf{z}_n\right)$. Under the same condition, the MLE is known to be Gaussian distributed [22]; that is, $\hat{\mathbf{v}}_{\text{MLE}} \sim \mathcal{N}(\mathbf{v}_o, \mathbf{G}^{-1})$.

Invertibility of $\mathbf{G}$ apparently depends on the non-zero entries of $\mathbf{a}$. As the number of PMU is small, especially at their initial deployment phase, state observability is at risk. By incorporating SCADA measurements, however, it is possible to regularize the system matrix, and thus enable state estimation even when $\mathbf{G}$ is singular [39]. Nonetheless, simply aggregating SCADA and PMU readings faces three challenges: 1) SCADA measurements are available every 4 s, whereas PMU ones are sampled every 0.03 s [13]; 2) explicitly including conventional measurements results in a nonlinear estimation problem of even higher dimensionality; and 3) upgrading the existing estimation software to accommodate high-rate PMU readings compromises backward compatibility [39]. An approach to address these challenges is through the so-called "pseudo-measurements," which are SCADA-based state estimates utilized as an extra set of measurements [31]. PMU readings and pseudo-measurements are then jointly used to form the MLE.

The alternative approach taken here treats state estimation in a Bayesian framework, where SCADA-based state estimates $\hat{\mathbf{v}}_s$ in rectangular coordinates are used as a priori information to aid PMU-based estimation. Specifically, it is assumed that the actual state vector is distributed as $\mathbf{v}_o \sim \mathcal{N}(\hat{\mathbf{v}}_s, \mathbf{\Sigma}_s)$—an assumption to be justified in Section IV-B, where an approximate expression for $\mathbf{\Sigma}_s$ is provided as well. Based on the linear model of (1) and the SCADA-based Gaussian prior information, the power system state is estimated as the vector yielding the maximum a posteriori probability (MAP) $p(\mathbf{v}|\{\mathbf{z}_n\}_{n=1}^{N_b})$. Because the latter is proportional to $p(\{\mathbf{z}_n\}_{n=1}^{N_b}|\mathbf{v})p(\mathbf{v})$, the MAP estimate $\hat{\mathbf{v}}_p$ can be easily shown to be the minimizer of

$$\min_{\mathbf{v}} \ \frac{1}{2}\sum_{n=1}^{N_b} a_n \|\mathbf{z}_n - \mathbf{H}_n\mathbf{v}\|^2_{\mathbf{\Sigma}_n^{-1}} + \frac{1}{2}\|\mathbf{v} - \hat{\mathbf{v}}_s\|^2_{\mathbf{\Sigma}_s^{-1}}. \quad (4)$$

The last term in (4) incorporates the SCADA-based prior state information and regularizes the MLE of (2). Note also that because of the linear-Gaussian model and the Gaussian prior, the MAP estimate coincides with the minimum mean-square error (MMSE) estimate [22].

The unique minimizer of (4) is provided straightforwardly in closed form as

$$\hat{\mathbf{v}}_p = \mathbf{G}_p^{-1}\left(\sum_{n=1}^{N_b} a_n \mathbf{H}_n^T \mathbf{\Sigma}_n^{-1}\mathbf{z}_n + \mathbf{\Sigma}_s^{-1}\hat{\mathbf{v}}_s\right) \quad (5)$$

where the now regularized gain matrix is

$$\mathbf{G}_p := \mathbf{G} + \mathbf{\Sigma}_s^{-1}. \quad (6)$$

Standard results from estimation theory assure that $\hat{\mathbf{v}}_p \sim \mathcal{N}(\mathbf{v}_o, \mathbf{G}_p^{-1})$ [22]. Thanks to the Gaussianity of $\hat{\mathbf{v}}_p$, matrix $\mathbf{G}_p$ is related to its η-confidence region, meaning that the estimate $\hat{\mathbf{v}}_p$ falls into the ellipsoid $(\hat{\mathbf{v}}_p - \mathbf{v}_o)^T\mathbf{G}_p(\hat{\mathbf{v}}_p - \mathbf{v}_o) \leq \alpha$ with probability η for $\alpha = F^{-1}_{\chi^2_{2N_b}}(\eta)$, where $F^{-1}_{\chi^2_{2N_b}}$ is the inverse cumulative distribution function of a chi-square random variable with $2N_b$ degrees of freedom [7, Sec. 7.5].

Phase Alignment: Since SCADA measurements are quadratically related to the state vector, complex bus voltages can be recovered only up to a phase rotation [28], [37]. This phase ambiguity is practically resolved by fixing the phase of a so-termed *reference bus*— typically enumerated as first—to zero. Equivalently, in rectangular coordinates, the entry $\mathrm{Im}(v_o^1)$ is set to zero, and is removed from the state vector. When state estimation is based solely on PMU measurements, phase ambiguity is not an issue, since measurements are linearly related to the system state. However, using the SCADA-based prior renders it necessary to "align" the prior with PMU phases. This is typically resolved by PMU-instrumenting the reference bus, and subtracting its measured voltage phase from all other PMU readings [31]. This clearly implies $a_1 = 1$, and again $\mathrm{Im}(v_o^1)$ is set to zero and removed from the state vector. For notational brevity, $\mathbf{v}_o$ will henceforth refer to the reduced $(2N_b - 1) \times 1$ vector, while $\mathbf{H}_n$'s ($\mathbf{G}$) will denote the corresponding matrices obtained after ignoring the $(N_b + 1)$st column (and row) from the respective matrices of (1) and (3). Moreover, the contribution $\mathbf{H}_1^T\mathbf{\Sigma}_1^{-1}\mathbf{H}_1$ of the reference bus to $\mathbf{G}$ in (3) can be added to the $\mathbf{\Sigma}_s^{-1}$ term when defining $\mathbf{G}_p$ in (6) without loss of generality. Non-reference buses are renumbered from one.

III. Optimal PMU Placement

Building on the state estimate in (4), the problem of PMU placement can be now stated as follows. Given

1) a power network of N_b buses (nodes);
2) matrices $\{\mathbf{H}_n, \mathbf{\Sigma}_n\}_{n=1}^{N_b}$ (c.f (1) and Section IV-A);
3) the covariance matrix $\mathbf{\Sigma}_s$ (cf. Section IV-B);
4) an integer $k \leq N_b$,

and assuming a PMU installed at the reference bus, the goal is to choose k buses to be PMU-instrumented so that the error of the estimator (5) is minimized.

PMU placement is cast here as a variation of the optimal experimental design problem [33]. The state estimation error for a specific PMU placement $\mathbf{a}$ has inverse covariance matrix $\mathbf{G}_p(\mathbf{a}) = \mathbf{G}(\mathbf{a}) + \mathbf{\Sigma}_s^{-1}$, where the dependence on $\mathbf{a}$ will be exemplified throughout this section.

Apparently, between two candidate placements $\mathbf{a}$ and $\mathbf{a}'$ with $\mathbf{1}_{N_b}^T\mathbf{a} = \mathbf{1}_{N_b}^T\mathbf{a}' = k$, placement $\mathbf{a}$ is preferable over $\mathbf{a}'$ if $\mathbf{G}_p(\mathbf{a}) \succeq \mathbf{G}_p(\mathbf{a}')$. However, the ordering over semi-positive matrices is partial: if $\mathbf{G}_p(\mathbf{a}) - \mathbf{G}_p(\mathbf{a}')$ is an indefinite matrix, i.e., it has both positive and negative eigenvalues, then none of the placements is better than the other. To overcome this ordering issue, placements are typically ranked based on a scalar-valued function of $\mathbf{G}_p^{-1}(\mathbf{a})$, $f(\mathbf{G}_p^{-1}(\mathbf{a}))$, that is to be minimized [7], [33]. Typical function choices are:

(c1) *E-optimal design*: $f_E(\mathbf{a}) := \lambda_{\max}(\mathbf{G}_p^{-1}(\mathbf{a}))$, where $\lambda_{\max}$ denotes the maximum eigenvalue of $\mathbf{G}_p^{-1}(\mathbf{a})$;

(c2) *A-optimal design*: $f_A(\mathbf{a}) := \mathrm{trace}(\mathbf{G}_p^{-1}(\mathbf{a}))$ that is equal to the sum of the eigenvalues of $\mathbf{G}_p^{-1}(\mathbf{a})$;

(c3) *M-optimal design*: $f_M(\mathbf{a}) := \max_i\{[\mathbf{G}_p^{-1}(\mathbf{a})]_{ii}\}$ corresponding to the maximum diagonal entry of $\mathbf{G}_p^{-1}(\mathbf{a})$; and

(c4) *D-optimal design*: $f_D(\mathbf{a}) := \log\det\left(\mathbf{G}_p^{-1}(\mathbf{a})\right)$, where $\det(\cdot)$ denotes matrix determinant.

Details and interesting geometric interpretations regarding choices (c1)–(c4) can be found in [7, Sec. 7.5] and [33]. After this scalarization step, the C-optimal PMU placement can be found as the solution of the optimization problem

$$\check{\mathbf{a}}_C := \arg\min_{\mathbf{a}} \quad f_C(\mathbf{a}) \quad (7a)$$

$$\text{s.t.} \quad \mathbf{a}^T\mathbf{1}_{N_b} = k \quad (7b)$$

$$\mathbf{a} \in \{0,1\}^{N_b} \quad (7c)$$

where $C \in \{E, A, M, D\}$. The constraint (7b) expresses the PMU budget. Trivially, when some buses have already been PMU-instrumented, they can be handled in the same way as for the reference bus (cf. end of Section III).

Unfortunately, solving (7) incurs combinatorial complexity [7], [21]. Suboptimal solutions can be obtained by converting the troublesome binary constraint (7c) to the box constraint $\mathbf{0}_{N_b} \leq \mathbf{a} \leq \mathbf{1}_{N_b}$ with the inequalities understood entry-wise [7]. Since $\{0,1\}^{N_b} \subset [0,1]^{N_b}$, such a conversion constitutes a *relaxation* of the original problem. Upon defining the relaxed feasible set

$$\mathcal{A} := \{\mathbf{a} : \mathbf{a}^T\mathbf{1}_{N_b} = k,\ \mathbf{0}_{N_b} \leq \mathbf{a} \leq \mathbf{1}_{N_b}\} \quad (8)$$

the relaxed problem is expressed as

$$\check{\mathbf{a}}_C := \arg\min_{\mathbf{a}\in\mathcal{A}} f_C(\mathbf{a}). \tag{9}$$

Fortunately, the problem in (9) can be efficiently solved as a convex optimization problem for all four cost function choices $C \in \{E, A, M, D\}$ [7], [21]. The SDP formulations of the problems are presented briefly next for completeness.

For the E-optimal design, instead of minimizing $\lambda_{\max}(\mathbf{G}_p^{-1}(\mathbf{a}))$, one can alternatively maximize $\lambda_{\min}(\mathbf{G}_p(\mathbf{a}))$ by introducing the auxiliary variable t and solving the SDP

$$\begin{aligned}(\check{\mathbf{a}}_E, \check{t}_E) := \arg\min_{\mathbf{a}\in\mathcal{A},t} \quad & -t \\ \text{s.t.} \quad & \mathbf{G}(\mathbf{a}) + \mathbf{\Sigma}_s^{-1} \succeq t \cdot \mathbf{I}_{2N_b-1}.\end{aligned} \tag{10}$$

For the A-optimal design, minimizing the trace of $\mathbf{G}_p^{-1}(\mathbf{a})$ can be accomplished after introducing the auxiliary vector variable $\mathbf{t} := [t_1 \ldots t_{2N_b-1}]^T$, leading to the SDP

$$\begin{aligned}(\check{\mathbf{a}}_A, \check{\mathbf{t}}_A) := \arg\min_{\mathbf{a}\in\mathcal{A},\mathbf{t}} \quad & \mathbf{t}^T \mathbf{1}_{2N_b-1} \\ \text{s.t.} \quad & \begin{bmatrix} \mathbf{G}(\mathbf{a})+\mathbf{\Sigma}_s^{-1} & \mathbf{e}_l \\ \mathbf{e}_l^T & t_l \end{bmatrix} \succeq \mathbf{0},\ l = 1, \ldots, 2N_b-1\end{aligned} \tag{11}$$

where $\mathbf{e}_l$ is the lth canonical vector. The M-optimal design can be suboptimally solved by the SDP

$$\begin{aligned}(\check{\mathbf{a}}_M, \check{t}_M) := \arg\min_{\mathbf{a}\in\mathcal{A},t} \quad & t \\ \text{s.t.} \quad & \begin{bmatrix} \mathbf{G}(\mathbf{a})+\mathbf{\Sigma}_s^{-1} & \mathbf{e}_l \\ \mathbf{e}_l^T & t \end{bmatrix} \succeq \mathbf{0},\ l = 1, \ldots, 2N_b-1.\end{aligned} \tag{12}$$

Finally, the relaxed D-optimal placement can be expressed as the convex optimization problem

$$\check{\mathbf{a}}_D := \arg\min_{\mathbf{a}\in\mathcal{A}} -\log\det\left(\mathbf{G}(\mathbf{a}) + \mathbf{\Sigma}_s^{-1}\right) \tag{13}$$

that can be efficiently solved via standard software [24], [35].

Even though (9) is solved offline whenever a new batch of acquired units is to be installed, interior point-based methods are deemed non-operational when N_b is larger than few hundreds [24], [35] (cf. Section V), in which case one has to resort to first-order optimization algorithms. To this end, a gradient projection algorithm for solving (9) for $C \in \{A, D\}$ is developed in the Appendix. For minimizing the convex, yet non-smooth functions $f_E(\mathbf{a})$ and $f_M(\mathbf{a})$ over $\mathcal{A}$, proximal gradient counterparts can be adopted, too.

The minimizers $\check{\mathbf{a}}_C$ of (10)–(13) do not necessarily have binary entries. A simple heuristic to obtain a binary solution is to set the k largest entries of $\check{\mathbf{a}}_C$ to 1, and zero the rest [7]. The so-acquired vector, denoted by $\hat{\mathbf{a}}_C$, belongs to the feasible set of the original non-convex problem (7), but in general it is not a minimizer of (7). It provides though the upper bound $f_C(\tilde{\mathbf{a}}_C) \leq f_C(\hat{\mathbf{a}}_C)$ for all C. Additionally, due to the relaxation, the minimizers of (10)–(13) yield also the lower bound $f_C(\check{\mathbf{a}}_C) \leq f_C(\tilde{\mathbf{a}}_C)$ for all placement criteria C. When the gap $f_C(\hat{\mathbf{a}}_C) - f_C(\check{\mathbf{a}}_C)$ becomes zero, the relaxation is deemed exact in the sense that $f_C(\hat{\mathbf{a}}_C) = f_C(\tilde{\mathbf{a}}_C)$ [21].

Extensions: The optimal PMU placement problem in (7) can accommodate several additional constraints imposed by communication, installation, or other practical considerations. Linear (in)equality constraints on the binary variables $\mathbf{a}$ can effect prior information of logical nature [21]. For example, the placement constraint that bus n can be PMU-instrumented only if bus m is, too, can be expressed as $a_m \geq a_n$; or, if a critical subnetwork $\mathcal{S}$ of the grid is to be equipped with at least k_s PMU devices, the constraint $\sum_{i\in\mathcal{S}} a_i \geq k_s$ should be added. Moreover, criteria other than state estimation error can be jointly considered in the placement process. For example, if having a PMU at the nth bus incurs communication cost γ_n, and there is a budget γ, the constraint $\sum_{n=1}^{N_b} \gamma_n a_n \leq \gamma$ can be attached to (7). All these constraints can be retained in the relaxed problems (9) by properly redefining its feasible set $\mathcal{A}$. However, the simple operation to acquire $\hat{\mathbf{a}}_C$ from $\check{\mathbf{a}}_C$ can no longer guarantee the feasibility of $\hat{\mathbf{a}}_C$.

Parameter Uncertainty: Problems (7) and (9) assume that matrices $\{\mathbf{H}_n, \mathbf{\Sigma}_n\}_{n=1}^{N_b}$ and $\mathbf{\Sigma}_s$ are perfectly known. However, measurement errors, seasonal variations, and the approximation involved in $\mathbf{\Sigma}_s$ introduce errors. Small deviations from nominal values can be lumped into the noise in (1), or can be handled possibly by following a total least-squares approach [20]. Furthermore, by postulating an uncertainty model, one can resort to robust (minimax) optimal experimental designs. Interestingly, under practical uncertainty models, the relaxed robust E-/D-optimal designs can be formulated as SDP problems, too [17], [21].

IV. Input Data

In this section, the input data for the optimal PMU placement problem are expressed in terms of physical properties of the power network. The regression matrices $\{\mathbf{H}_n\}$ appearing in (1) are described in Section IV-A, while an approximation of $\mathbf{\Sigma}_s$ required in (4) is provided in Section IV-B.

A. Regression Matrices

Let $\tilde{\mathbf{v}} = \mathbf{v}_r + j\mathbf{v}_i$ be the $N_b \times 1$ vector of complex nodal voltages. The vector of complex currents injected per bus is

$$\tilde{\mathbf{i}} = \mathbf{Y}\tilde{\mathbf{v}} \tag{14}$$

where $\mathbf{Y} \in \mathbb{C}^{N_b\times N_b}$ denotes the bus admittance matrix, and can be explicitly expressed as a function of the bus admittances, the line series admittances and charging susceptances, and potential tap ratios and phase shifters [41]. Considering a power network consisting of N_l lines, the $2N_l \times 1$ vector of complex currents injected on each line can be described similarly by

$$\tilde{\mathbf{i}}_{\text{fl}} = \mathbf{Y}_{\text{fl}}\tilde{\mathbf{v}} \tag{15}$$

where $\mathbf{Y}_{\text{fl}} \in \mathbb{C}^{2N_l\times N_b}$ denotes the line-to-bus admittance matrix [41]. Note that the current flowing from bus m to bus n is not equal to the negative of the current flowing in the reverse direction. Hence, each line is considered twice in vector $\tilde{\mathbf{i}}_{\text{fl}}$. The regression matrices $\mathbf{H}_n$ in (1) can now be expressed as

$$\mathbf{H}_n = \begin{bmatrix} \mathbf{e}_n^T & \mathbf{0}^T \\ \mathbf{0}^T & \mathbf{e}_n^T \\ \mathbf{S}_n\text{Re}(\mathbf{Y}_{\text{fl}}) & -\mathbf{S}_n\text{Im}(\mathbf{Y}_{\text{fl}}) \\ \mathbf{S}_n\text{Im}(\mathbf{Y}_{\text{fl}}) & \mathbf{S}_n\text{Re}(\mathbf{Y}_{\text{fl}}) \end{bmatrix} \tag{16}$$

where $\mathbf{S}_n$ is the binary $L_n \times 2N_l$ matrix selecting the rows of $\mathbf{Y}_{\text{fl}}$ corresponding to the lines originating from bus n.

B. Covariance of the SCADA-Based State Estimate

Conventional power meters measure subsets of nodal real and reactive power injections, real and reactive power line flows, as well as nodal voltage magnitudes. Consider M such measurements concatenated in $\mathbf{z}_s \in \mathbb{R}^M$ that are related to $\mathbf{v}_o$ through the generally nonlinear function $\mathbf{h}(\cdot) : \mathbb{R}^{2N_b} \to \mathbb{R}^M$. The SCADA reading model is

$$\mathbf{z}_s = \mathbf{h}(\mathbf{v}_o) + \mathbf{w}_s \tag{17}$$

where $\mathbf{w}_s \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_e)$ denotes the noise vector. Supposing that the SCADA-based estimate $\hat{\mathbf{v}}_s$ has converged to the MLE based on (17), the estimate $\hat{\mathbf{v}}_s$ is asymptotically (as $M \to \infty$) normal; that is, $\hat{\mathbf{v}}_s \overset{a}{\sim} \mathcal{N}(\mathbf{v}_o, \mathbf{\Sigma}_s(\mathbf{v}_o))$, where [22]

$$\mathbf{\Sigma}_s(\mathbf{v}_o) = \left(\mathbf{J}^T(\mathbf{v}_o)\mathbf{\Sigma}_e^{-1}\mathbf{J}(\mathbf{v}_o)\right)^{-1} \tag{18}$$

and $\mathbf{J}(\mathbf{v}_o) := \nabla_{\mathbf{v}}\mathbf{h}(\mathbf{v})|_{\mathbf{v}=\mathbf{v}_o}$ is the $M \times 2N_b$ Jacobian matrix of $\mathbf{h}(\mathbf{v})$ evaluated at $\mathbf{v} = \mathbf{v}_o$. Since $\mathbf{h}(\mathbf{v})$ is nonlinear, the covariance matrix in (18) depends on the actual nodal voltage values $\mathbf{v}_o$, which are unknown. To resolve this vicious cycle, one approach is to surrogate (18) by $\mathbf{\Sigma}_s(\hat{\mathbf{v}}_s)$ in (5).

For optimal PMU placement, however, the goal is to solve (7) or (9) without having acquired any measurements and for all possible values of the state vector $\mathbf{v}_o$. Toward this end, the idea here is to replace $\mathbf{v}_o$ in (18) by the so-called flat voltage profile $\mathbf{v}_{\text{flat}} := [\mathbf{1}_{N_b}^T \ \mathbf{0}_{N_b}^T]^T$, commonly used for initializing state estimation algorithms [37]. For this reason, wherever $\mathbf{\Sigma}_s$ appears in the previous section, it is replaced by $\mathbf{\Sigma}_s(\mathbf{v}_{\text{flat}})$. Then, the Jacobian $\mathbf{J}(\mathbf{v}_{\text{flat}})$ can be neatly expressed in terms of $\mathbf{Y}$ and $\mathbf{Y}_{\text{fl}}$. Note that other typical voltage profiles can be alternatively used, too.

What remains to be determined is the Jacobian matrix $\mathbf{J}(\mathbf{v}_{\text{flat}})$ for the different types of SCADA measurements. Starting with voltage magnitudes, if a row of $\mathbf{J}(\mathbf{v}_{\text{flat}})$ corresponds to the voltage magnitude measurement at bus n, namely $\sqrt{v_{r,n}^2 + v_{i,n}^2}$, then it is readily given by $[\mathbf{e}_n^T \ \mathbf{0}^T]$. Before proceeding with power SCADA measurements, note that because $\tilde{\mathbf{v}} = \mathbf{v}_r + j\mathbf{v}_i$, it holds that [cf.(14)]

$$\frac{\partial \tilde{\mathbf{v}}}{\partial \mathbf{v}_r} = \mathbf{I}_{N_b}, \quad \frac{\partial \tilde{\mathbf{v}}}{\partial \mathbf{v}_i} = j\mathbf{I}_{N_b} \tag{19a}$$

$$\frac{\partial \tilde{\mathbf{i}}}{\partial \mathbf{v}_r} = \mathbf{Y}, \quad \frac{\partial \tilde{\mathbf{i}}}{\partial \mathbf{v}_i} = j\mathbf{Y}. \tag{19b}$$

The vector of complex powers injected at each bus $\mathbf{s} \in \mathbb{C}^{N_b}$ is provided by the formula $\mathbf{s} = \text{diag}(\tilde{\mathbf{v}})\tilde{\mathbf{i}}^* = \text{diag}(\tilde{\mathbf{i}}^*)\tilde{\mathbf{v}}$. After applying the product and the chain rules of differentiation, and using (19), the power gradients can be expressed as

$$\frac{\partial \mathbf{s}}{\partial \mathbf{v}_r} = \text{diag}(\tilde{\mathbf{v}})\mathbf{Y}^* + \text{diag}(\mathbf{Y}^*\tilde{\mathbf{v}}^*) \tag{20a}$$

$$\frac{\partial \mathbf{s}}{\partial \mathbf{v}_i} = -j\text{diag}(\tilde{\mathbf{v}})\mathbf{Y}^* + j\text{diag}(\mathbf{Y}^*\tilde{\mathbf{v}}^*). \tag{20b}$$

When these gradients are evaluated at $\mathbf{v}_{\text{flat}}$, they yield

$$\left.\frac{\partial \mathbf{s}}{\partial \mathbf{v}_r}\right|_{\mathbf{v}_{\text{flat}}} = \mathbf{Y}^* + \text{diag}(\mathbf{Y}^*\mathbf{1}_{N_b}) \tag{21a}$$

$$\left.\frac{\partial \mathbf{s}}{\partial \mathbf{v}_i}\right|_{\mathbf{v}_{\text{flat}}} = -j\mathbf{Y}^* + j\text{diag}(\mathbf{Y}^*\mathbf{1}_{N_b}). \tag{21b}$$

Let $\mathbf{p}, \ \mathbf{q} \in \mathbb{R}^{N_b}$ denote the vectors of real and reactive nodal power injections, respectively. Since $\mathbf{s} = \mathbf{p} + j\mathbf{q}$ and after using (21), the gradient evaluations are

$$\left.\frac{\partial \mathbf{p}}{\partial \mathbf{v}_r}\right|_{\mathbf{v}_{\text{flat}}} = \text{Re}(\mathbf{Y}) + \text{diag}(\text{Re}(\mathbf{Y})\mathbf{1}_{N_b}) \tag{22a}$$

$$\left.\frac{\partial \mathbf{p}}{\partial \mathbf{v}_i}\right|_{\mathbf{v}_{\text{flat}}} = -\text{Im}(\mathbf{Y}) + \text{diag}(\text{Im}(\mathbf{Y})\mathbf{1}_{N_b}) \tag{22b}$$

$$\left.\frac{\partial \mathbf{q}}{\partial \mathbf{v}_r}\right|_{\mathbf{v}_{\text{flat}}} = -\text{Im}(\mathbf{Y}) - \text{diag}(\text{Im}(\mathbf{Y})\mathbf{1}_{N_b}) \tag{22c}$$

$$\left.\frac{\partial \mathbf{q}}{\partial \mathbf{v}_i}\right|_{\mathbf{v}_{\text{flat}}} = -\text{Re}(\mathbf{Y}) + \text{diag}(\text{Re}(\mathbf{Y})\mathbf{1}_{N_b}). \tag{22d}$$

Hence, if a row of $\mathbf{J}(\mathbf{v}_{\text{flat}})$ corresponds to the real/reactive power injected at bus n, then it is provided by the nth row of (22a)/(22c) followed by the nth row of (22b)/(22d).

Similar derivations can be followed for the line power flows. More precisely, the complex power flowing from bus m to bus n through the lth line (m, n) is $\tilde{v}_m \tilde{i}_{\text{fl},l}^*$, where $\tilde{v}_m$ is the mth entry of $\tilde{\mathbf{v}}$ and $\tilde{i}_{\text{fl},l}$ is the lth entry of $\tilde{\mathbf{i}}_{\text{fl}}$ in (15). Next, define the line-to-source bus adjacency matrix $\mathbf{S} \in \mathbb{R}^{2N_l \times N_b}$ such that if its lth row corresponds to the directed (m, n) line, the mth element of this row is 1, and the rest are zero. Let also $\mathbf{p}_{\text{fl}}, \ \mathbf{q}_{\text{fl}} \in \mathbb{R}^{2N_l}$ denote the vectors of real and reactive line power injections, respectively, following the line ordering of (15). Then, the line power injections can be neatly expressed by the matrix-vector formula

$$\mathbf{p}_{\text{fl}} + j\mathbf{q}_{\text{fl}} = \text{diag}(\mathbf{S}\tilde{\mathbf{v}})\mathbf{Y}_{\text{fl}}^*\tilde{\mathbf{v}}^* = \text{diag}(\mathbf{Y}_{\text{fl}}^*\tilde{\mathbf{v}}^*)\mathbf{S}\tilde{\mathbf{v}}. \tag{23}$$

Based on (23), mimicking the development of (19)–(22), and exploiting the fact that $\mathbf{S}\mathbf{1}_{N_b} = \mathbf{1}_{2N_l}$, it follows readily that

$$\left.\frac{\partial \mathbf{p}_{\text{fl}}}{\partial \mathbf{v}_r}\right|_{\mathbf{v}_{\text{flat}}} = \text{Re}(\mathbf{Y}_{\text{fl}}) + \text{diag}(\text{Re}(\mathbf{Y}_{\text{fl}})\mathbf{1}_{N_b})\mathbf{S} \tag{24a}$$

$$\left.\frac{\partial \mathbf{p}_{\text{fl}}}{\partial \mathbf{v}_i}\right|_{\mathbf{v}_{\text{flat}}} = -\text{Im}(\mathbf{Y}_{\text{fl}}) + \text{diag}(\text{Im}(\mathbf{Y}_{\text{fl}})\mathbf{1}_{N_b})\mathbf{S} \tag{24b}$$

$$\left.\frac{\partial \mathbf{q}_{\text{fl}}}{\partial \mathbf{v}_r}\right|_{\mathbf{v}_{\text{flat}}} = -\text{Im}(\mathbf{Y}_{\text{fl}}) - \text{diag}(\text{Im}(\mathbf{Y}_{\text{fl}})\mathbf{1}_{N_b})\mathbf{S} \tag{24c}$$

$$\left.\frac{\partial \mathbf{q}_{\text{fl}}}{\partial \mathbf{v}_i}\right|_{\mathbf{v}_{\text{flat}}} = -\text{Re}(\mathbf{Y}_{\text{fl}}) + \text{diag}(\text{Re}(\mathbf{Y}_{\text{fl}})\mathbf{1}_{N_b})\mathbf{S}. \tag{24d}$$

Likewise for (22), if a row of $\mathbf{J}(\mathbf{v}_{\text{flat}})$ corresponds to the real/reactive power flowing over the directed line l, then it is provided by the lth row of (24a)/(24c) followed by the lth row of (24b)/(24d).

V. Simulated Tests

The relaxed PMU placement methods are evaluated numerically using MATLAB [26] on an Intel Duo Core @ 2.2 GHz

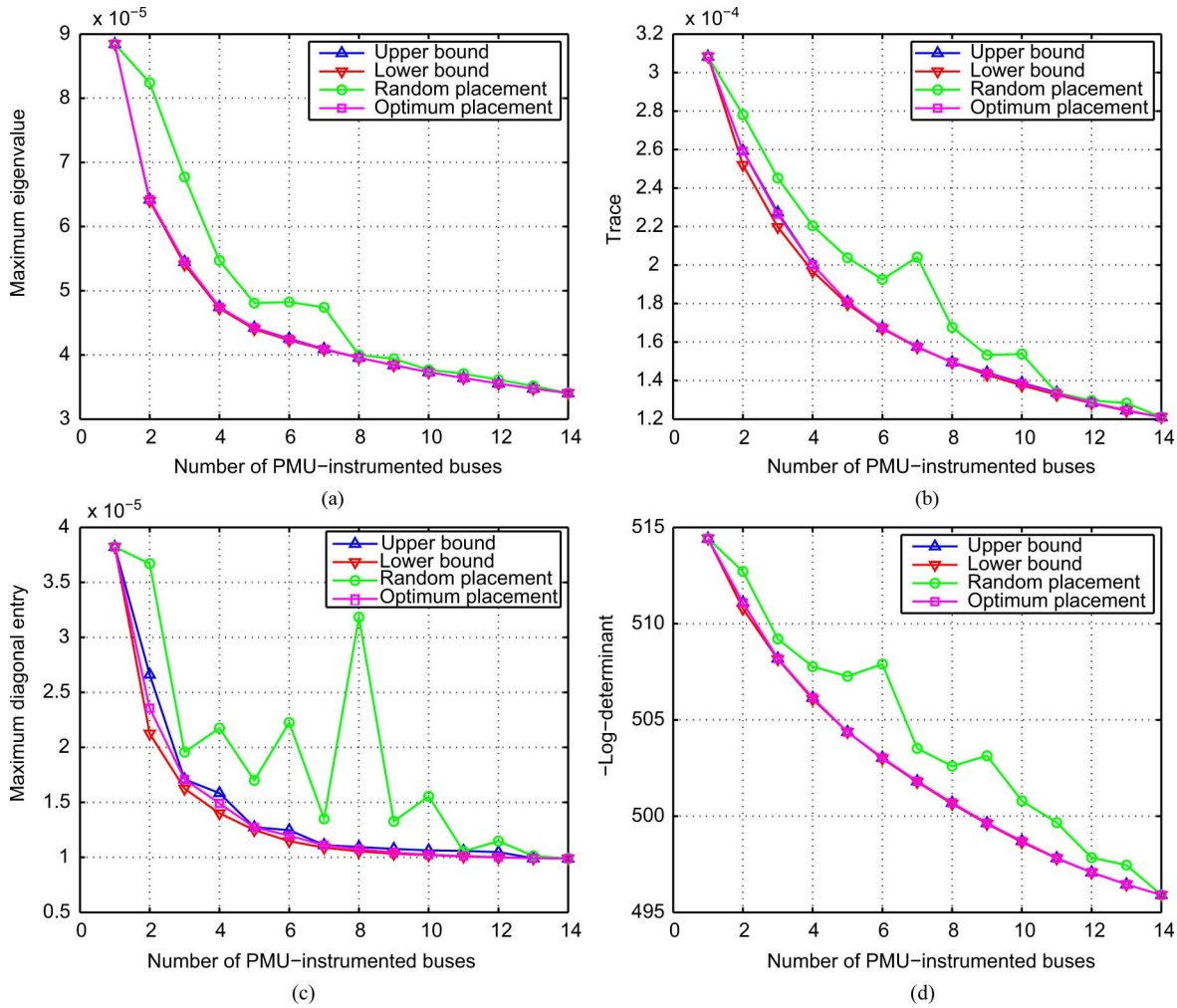


Fig. 1. IEEE 14-bus test case: $f_C(\mathbf{a})$ attained by solving (7) and (9), and by a random PMU placement for the different criteria $C \in \{E, A, M, D\}$. (a) E-optimal design. (b) A-optimal design. (c) M-optimal design. (d) D-optimal design.

(4GB RAM) computer. Three commonly used power network benchmarks, namely the IEEE 14-, 30-, and 118-bus systems, are considered [36]. The bus- and line-to-bus admittance matrices $\mathbf{Y}$ and $\mathbf{Y}_{\text{fl}}$, respectively, are obtained using the MATPOWER software [41], while the involved optimization parameters are stored as sparse matrices.

Regarding SCADA measurements, the redundancy ratio, which is the ratio of SCADA measurements over the number of state variables, is typically around 2.2 [28]. Given that the (N_b, N_l) pairs of the three power networks considered here are (14,20), (30,41), and (118,186), this ratio was approximately obtained by measuring the 50% of bus voltage magnitudes and (real/reactive) powers at buses and lines. The locations of SCADA measurements are initially selected uniformly at random and remain fixed throughout the experiments.

The covariance matrix of the SCADA measurement noise $\mathbf{\Sigma}_e$ [c.f (17)] is modeled as diagonal: the standard deviation for voltage magnitude, bus power injections, and line power flows is 0.01, 0.015, and 0.02, respectively, in accordance with the state estimation setups in [41]. PMU measurements are assumed to have a diagonal covariance matrix $\mathbf{\Sigma}_n$ [c.f (1)], too. Compliant with SCADA readings, the standard deviation for bus voltage and line current PMU measurements is 0.01 and 0.02, respectively. Note that the proposed PMU deployment method applies equally well even for non-diagonal $\mathbf{\Sigma}_n$'s.

Starting with the IEEE 14-bus test case, the four placement criteria were evaluated for $k \in [1, 14]$ units. Fig. 1 shows the cost values for: 1)–2) $f_C(\hat{\mathbf{a}}_C)$ and $f_C(\check{\mathbf{a}}_C)$ acquired by solving (9) with the SDPT3 and YALMIP softwares [24], [35] in a matter of a few seconds; 3) a random PMU placement; and 4) the cost $f_C(\tilde{\mathbf{a}}_C)$ obtained by exhaustively solving (7); all for $C \in \{E, A, M, D\}$. In most cases tested, $\hat{\mathbf{a}}_C$ coincided with $\tilde{\mathbf{a}}_C$, as verified by the zero gap $f(\hat{\mathbf{a}}_C) - f(\check{\mathbf{a}}_C)$. In general, $\hat{\mathbf{a}}_M$ differs from $\tilde{\mathbf{a}}_M$ by at most one unit.

Regarding the 14-bus test case, note first that the minimizers $\tilde{\mathbf{a}}_E$ and $\tilde{\mathbf{a}}_M$ do not possess a nesting property for varying k, meaning that the optimal k-placement is not necessarily a subset of the optimal $(k+1)$-one. Secondly, the numerical tests verify the importance of the SCADA-based prior: For the A-optimal placement and $k = 14$, the trace of $\mathbf{G}_p^{-1}(\tilde{\mathbf{a}}_A)$ increases from 1×10^{-4} to 2.5×10^{-4} when the prior is ignored. More critically, without the SCADA-based prior in (6), one needs $k \geq 5$ units to obtain a non-singular $\mathbf{G}_p(\tilde{\mathbf{a}}_A)$.

Next, the effect of PMU placement on state estimation accuracy is investigated. Fig. 2 illustrates the standard deviation of the real (top) and the imaginary (bottom) parts of nodal voltages

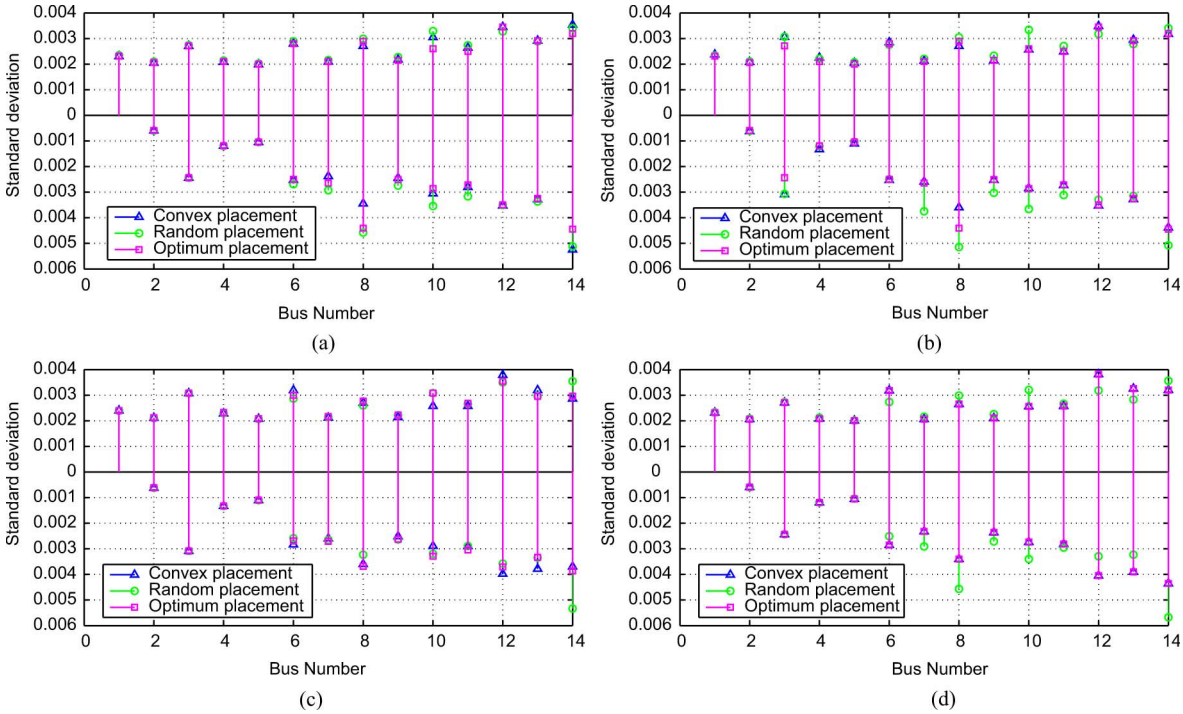


Fig. 2. IEEE 14-bus test case: Standard deviation for the real (top) and imaginary (bottom) nodal voltages for $k = 4$ PMU. (a) E-optimal design. (b) A-optimal design. (c) M-optimal design. (d) D-optimal design.

after deploying $k = 4$ phasor measurement units according to the placements $\hat{\mathbf{a}}_C$, $\tilde{\mathbf{a}}_C$, and a random one. The performance of the $\hat{\mathbf{a}}_C$ placement is close to that of $\tilde{\mathbf{a}}_C$ for all four criteria.

For the IEEE 30-bus benchmark, solving (7) for an arbitrary $k \in [1, 30]$ is prohibitively complex; hence, Fig. 3 depicts only the costs attained by solving (9) using SDPT3/YALMIP [24], [35], within a few seconds; and by a random placement. Even though PMU misplacements compared to $\tilde{\mathbf{a}}_C$ cannot be quantified, the small gap $f_C(\hat{\mathbf{a}}_C) - f_C(\breve{\mathbf{a}}_C)$ indicates the closeness of the obtained solutions to the optimum ones.

Finally, the IEEE 118-bus test case is considered. The network size disqualifies any combinatorial search. SDPT3/YALMIP can solve (9) for $C \in \{E, D\}$ in a few seconds, but the cases of $\{A, M\}$ cannot be handled. The relaxed A-optimal deployment is solved using the gradient projection algorithm developed in the Appendix. Depending on the value of k, the algorithm converges in 100–200 iterations, which is always less than 1.5 min. Fig. 4 shows the quite impressive results obtained by the convex relaxations for $C \in \{E, A\}$; similar results hold for $C = D$, but are omitted due to space limitations.

VI. Conclusions

Recognizing that system state accuracy is a key factor in enabling situational awareness for the power grid, PMU deployment was approached here based on estimation theoretic criteria. The estimation of nodal voltages expressed in rectangular coordinates by using both voltage and current PMU measurements was formulated as a linear regression problem, while SCADA-derived state estimates were incorporated in the form of Gaussian priors. By posing the problem under the framework of optimal experimental design, the involved error covariance matrix was then minimized in a well-defined sense. To obtain practically computable solutions, the combinatorial problems involved were relaxed to convex optimization ones. Surprisingly, numerical tests using standard SDP solvers and a devised gradient projection algorithm demonstrated that the obtained PMU placements come close to the optimum ones as verified by the small or oftentimes zero suboptimality gap for the tested IEEE bus cases.

Appendix
A Gradient Projection Algorithm

Problem (9) for $C \in \{A, D\}$ entails minimization of a convex, continuously differentiable function over $\mathcal{A}$. A gradient projection algorithm for accomplishing this task is presented here. The algorithm comprises the iterative updates

$$\mathbf{a}_C^{r+1} = [\mathbf{a}_C^r - s_r \mathbf{g}_C(\mathbf{a}_C^r)]_{\mathcal{A}} \tag{25}$$

where s_r denotes the step size; $\mathbf{g}_C(\mathbf{a}) := \nabla f_C(\mathbf{a})$; and the symbol $[\mathbf{c}]_{\mathcal{A}}$ denotes the projection operator onto $\mathcal{A}$, i.e.,

$$[\mathbf{c}]_{\mathcal{A}} := \arg\min_{\mathbf{z}} \left\{ \|\mathbf{z} - \mathbf{c}\|_2^2 : \mathbf{z} \in \mathcal{A} \right\}. \tag{26}$$

The ingredients of the algorithm are detailed next.

Gradient Vectors: Using matrix calculus [8], the nth entry of $\mathbf{g}_D(\mathbf{a})$ for $n = 1, \ldots, N_b$ can be expressed as

$$[\mathbf{g}_D(\mathbf{a})]_n = -\operatorname{trace}\left(\mathbf{G}_p^{-1}(\mathbf{a}) \cdot \mathbf{H}_n^T \mathbf{\Sigma}_n^{-1} \mathbf{H}_n\right) \tag{27a}$$

$$= -\operatorname{trace}\left(\tilde{\mathbf{H}}_n \mathbf{G}_p^{-1}(\mathbf{a}) \tilde{\mathbf{H}}_n^T\right) \tag{27b}$$

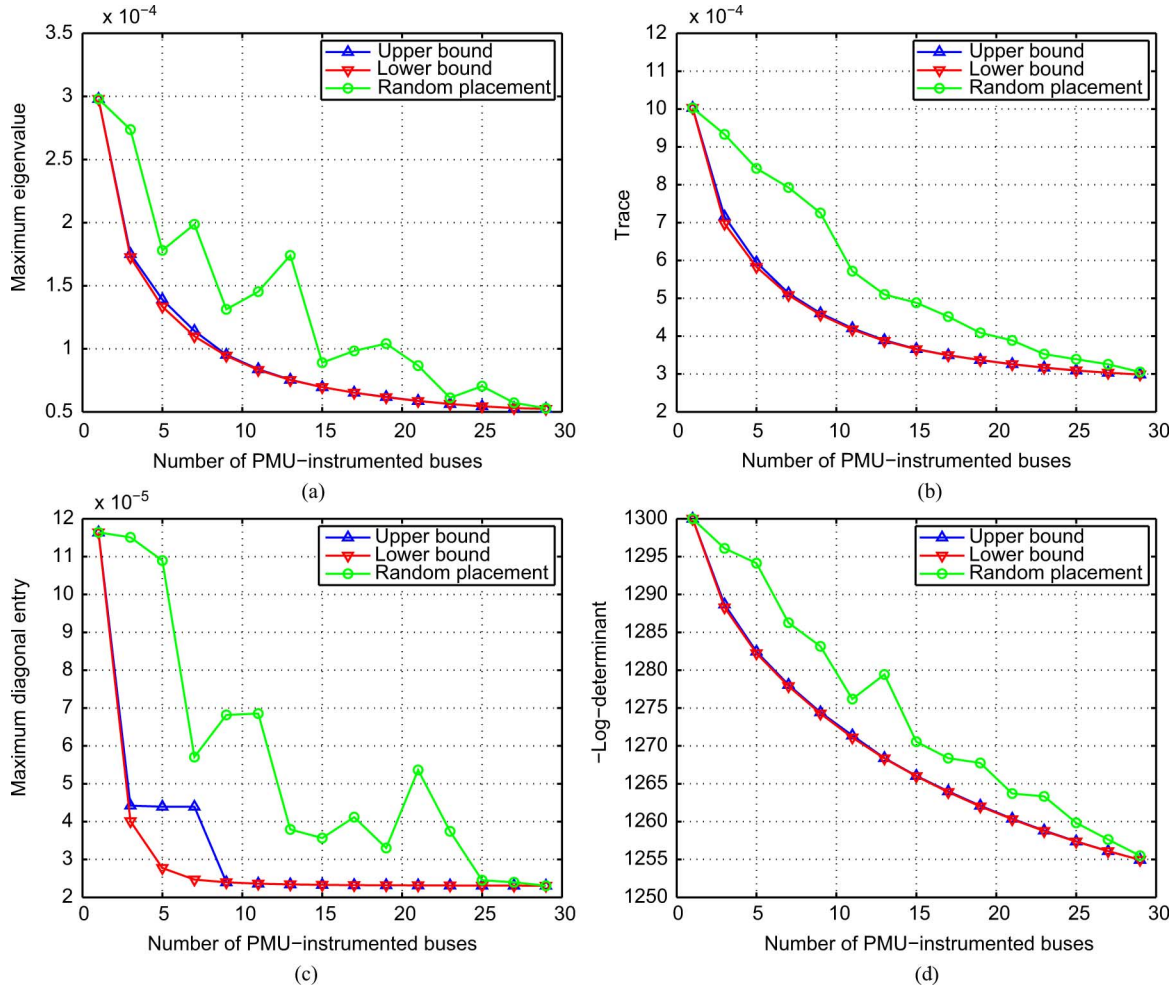


Fig. 3. IEEE 30-bus test case: $f_C(\mathbf{a})$ attained by solving (9) and by a random PMU placement for the different criteria $C \in \{E, A, M, D\}$. (a) E-optimal design. (b) A-optimal design (c) M-optimal design. (d) D-optimal design.

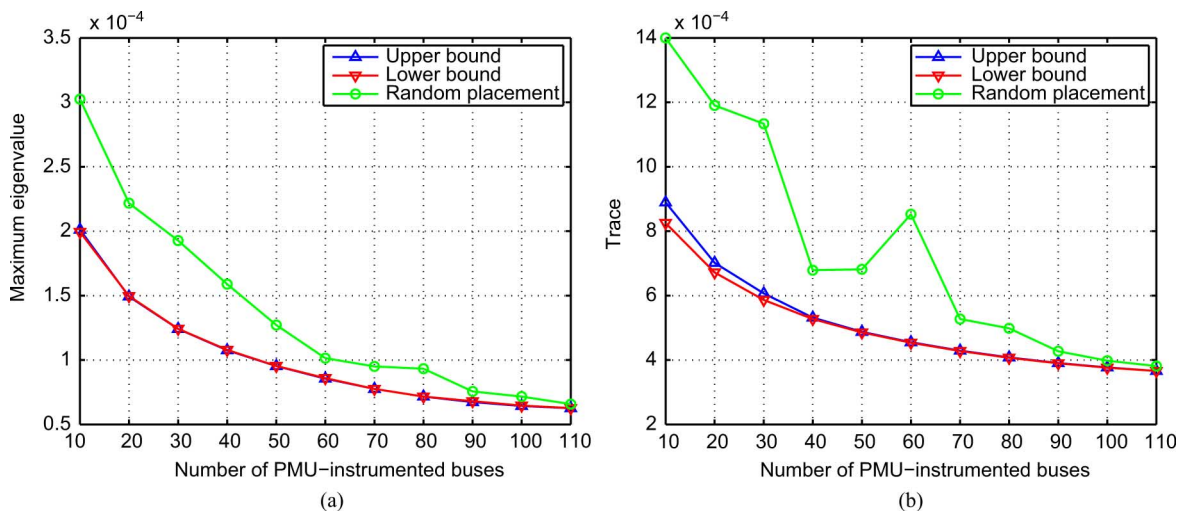


Fig. 4. IEEE 118-bus test case: $f_C(\mathbf{a})$ attained by solving (9) and by a random PMU placement for the different criteria $C \in \{E, A, M, D\}$. (a) E-optimal design. (b) A-optimal design.

where $\tilde{\mathbf{H}}_n := \mathbf{\Sigma}^{-1/2}\mathbf{H}_n$; and likewise for $\mathbf{g}_A(\mathbf{a})$

$$[\mathbf{g}_A(\mathbf{a})]_n = -\operatorname{trace}\left(\mathbf{G}_p^{-2}(\mathbf{a}) \cdot \mathbf{H}_n^T \mathbf{\Sigma}_n^{-1} \mathbf{H}_n\right) \quad (28a)$$

$$= -\operatorname{trace}\left(\mathbf{G}_n^T(\mathbf{a})\mathbf{G}_n(\mathbf{a})\right) \quad (28b)$$

where $\mathbf{G}_n(\mathbf{a}) := \mathbf{G}_p^{-1}(\mathbf{a})\tilde{\mathbf{H}}_n^T$. Efficient implementation of (27) and (28) leverages three facts: 1) the involved matrices have sparse structure; 2) the formulas (27b) and (28b) are computationally preferable over (27a) and (28a), since $\mathbf{H}_n$'s are $M_n \times$

$(2N_b - 1)$ with $M_n \ll N_b$; and 3) inverting $\mathbf{G}_p(\mathbf{a})$ can be accelerated by successively applying the matrix inversion lemma on (3) and (6) [8].

Projection onto $\mathcal{A}$: The Karush-Kuhn-Tucker (KKT) conditions for problem (26) show that its unique minimizer $[\mathbf{c}]_{\mathcal{A}}$ is provided by the thresholding rule for $n = 1, \ldots, N_b$

$$[\mathbf{z}(\lambda)]_n := \begin{cases} 0, & c_n \leq \lambda \\ c_n - \lambda, & \lambda < c_n < \lambda + 1 \\ 1, & c_n \geq \lambda + 1 \end{cases} \tag{29}$$

and for a value of λ, named $\lambda^\star$, that satisfies the constraint $\mathbf{1}_{N_b}^T \mathbf{z}(\lambda^\star) = k$. It can be seen that $\mathbf{1}_{N_b}^T \mathbf{z}(\lambda)$ is a continuous decreasing function of λ; hence, $\lambda^\star$ and $[\mathbf{c}]_{\mathcal{A}} = \mathbf{z}(\lambda^\star)$ can be easily found via bisection as shown in the tabulated Algorithm 1 .

The algorithm is initialized at $\mathbf{a}_C^0 = (k/N_b)\mathbf{1}_{N_b}$. At iteration r, vector $\mathbf{g}_C(\mathbf{a}_C^r)$ is found using (27) and (28); the step size s_r can be fixed to a sufficiently small value (see e.g., [6] for details); and the projection required in (25) is performed using Algorithm 1 for $\mathbf{c} = \mathbf{a}_C^r - s_r \mathbf{g}_C(\mathbf{a}_C^r)$. The algorithm can be terminated when $\left\|\mathbf{a}_C^{r+1} - \mathbf{a}_C^r\right\|_2 \leq 10^{-4}$; and its convergence is guaranteed [6].

Algorithm 1: Projection Onto $\mathcal{A}$

Require: Vector $\mathbf{c}$ and scalar $k \geq 0$.

1: Initialize $\lambda_{\min} = \min_n c_n - 1$, and $\lambda_{\max} = \max_n c_n$.

2: **repeat**

3: Update $\lambda = (\lambda_{\min} + \lambda_{\max})/2$.

4: Find $\mathbf{z}(\lambda)$ using **(29)**.

5: **If** $\mathbf{1}_{N_b}^T \mathbf{z}(\lambda) > k$, set $\lambda_{\min} = \lambda$;

6: **else** set $\lambda_{\max} = \lambda$.

7: **until** $|\mathbf{1}_{N_b}^T \mathbf{z}(\lambda) - k| \leq 10^{-4}$.

REFERENCES

[1] N. H. Abbasy and H. M. Ismail, "A unified approach for the optimal PMU location for power system state estimation," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 806–813, May 2009.

[2] F. Aminifar, M. Fotuhi-Firuzabad, M. Shahidehpour, and A. Khodaei, "Probabilistic multistage PMU placement in electric power systems," *IEEE Trans. Power Del.*, vol. 26, no. 2, pp. 841–849, Apr. 2011.

[3] F. Aminifar, C. Lucas, A. Khodaei, and M. Fotuhi-Firuzabad, "Optimal placement of phasor measurement units using immunity genetic algorithm," *IEEE Trans. Power Del.*, vol. 24, no. 3, pp. 1014–1020, Jul. 2009.

[4] M. Asprou and M. Kyriakides, "Optimal PMU placement for improving hybrid state estimator accuracy," in *Proc. Power Tech*, Trondheim, Norway, Jun. 2011.

[5] T. Baldwin, L. Mili, M. B. Boisen, and R. Adapa, "Power system observability with minimal phasor measurement placement," *IEEE Trans. Power Syst.*, vol. 8, no. 2, pp. 707–715, May 1993.

[6] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA: Athena Scientific, 1999.

[7] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York: Cambridge Univ. Press, 2004.

[8] M. Brooks, The Matrix Reference Manual, 2011. [Online]. Available: http://www.ee.imperial.ac.uk/hp/staff/dmb/matrix/intro.html.

[9] J. Y. Cai, Z. Huang, J. Hauer, and K. Martin, "Current status and experience of WAMS implementation in North America," in *Proc. IEEE Transmission and Distribution Conf.*, 2005, pp. 1–7.

[10] S. Chakrabarti and E. Kyriakides, "Optimal placement of phasor measurement units for power system observability," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1433–1440, Aug. 2008.

[11] S. Chakrabarti, E. Kyriakides, and D. Eliades, "Placement of synchronized measurements for power system observability," *IEEE Trans. Power Del.*, vol. 24, no. 1, pp. 12–19, Jan. 2009.

[12] K. A. Clements, "Observability methods and optimal meter placement," *Int. J. Elect. Power Energy Syst.*, vol. 12, no. 2, pp. 88–93, 1990.

[13] J. E. Dagle, "The North American synchrophasor initiative (NASPI)," in *Proc. IEEE PES General Meeting*, Detroit, MI, Jul. 2010.

[14] J. De La Ree, V. Centeno, J. Thorp, and A. Phadke, "Synchronized phasor measurement applications in power systems," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 20–27, Jun. 2010.

[15] D. Dua, S. Dambhare, R. K. Gajbhiye, and S. A. Soman, "Optimal multistage scheduling of PMU placement: An ILP approach," *IEEE Trans. Power Del.*, vol. 23, no. 4, pp. 1812–1820, Oct. 2006.

[16] R. Emami and A. Abur, "Robust measurement design by placing synchronized phasor measurements on network branches," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 38–43, Feb. 2010.

[17] P. Flaherty, M. Jordan, and A. Arkin, "Robust design of biological experiments," *Advances in Neural Information Processing Systems (NIPS)*, vol. 19, 2006.

[18] A. Z. Gamm, I. N. Kolosok, A. M. Glazunova, and E. S. Korkina, "PMU placement criteria for EPS state estimation," *Electric Utility Deregulation and Restructuring and Power Technologies*, Apr. 2008.

[19] B. Gou, "Optimal placement of PMUs by integer linear programming," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1525–1526, Aug. 2008.

[20] S. V. Huffel and J. Vandewalle, *The Total Least Squares Problem: Computational Aspects and Analysis*, ser. Frontier in Applied Mathematics. Philadelphia, PA: SIAM, 1991, vol. 9.

[21] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 451–462, Feb. 2009.

[22] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ: Prentice Hall, 1996.

[23] Q. Li, R. Negi, and M. D. Ilic, "Phasor measurement units placement for power system state estimation: A greedy approach," in *Proc. IEEE PES General Meeting*, Detroit, MI, Jul. 2011.

[24] J. Lofberg, "A toolbox for modeling and optimization in MATLAB," in *Proc. CACSD Conf.*, 2004. [Online]. Available: http://users.isy.liu.se/johanl/yalmip.

[25] V. Madani, M. Parashar, J. Giri, S. Durbha, F. Rahmatian, D. Day, M. Adamiak, and G. Sheble, "PMU placement considerations—a roadmap for optimal PMU placement," in *Proc. Power Systems Conf. Expo. (PSCE)*, Phoenix, AZ, Mar. 2011.

[26] *Matlab. Version 7.5.0*. Natick, MA: MathWorks, 2007.

[27] B. Milosevic and M. Begovic, "Nondominated sorting genetic algorithm for optimal phasor measurement placement," *IEEE Trans. Power Syst.*, vol. 18, no. 1, pp. 69–75, Feb. 2003.

[28] A. Monticelli, "Electric power system state estimation," *Proc. IEEE*, vol. 88, no. 2, pp. 262–282, Feb. 2000.

[29] North American SyncroPhasor Initiative, Synchrophasor Technology Roadmap, 2009.

[30] R. F. Nuqui and A. G. Phadke, "Phasor measurement unit placement techniques for complete and incomplete observability," *IEEE Trans. Power Del.*, vol. 20, no. 4, pp. 2381–2388, Oct. 2005.

[31] A. G. Phadke and J. S. Thorp, *Synchronized Phasor Measurements and Their Applications*. New York: Springer, 2008.

[32] A. G. Phadke, J. S. Thorp, and K. J. Karimi, "State estimation with phasor measurements," *IEEE Trans. Power Syst.*, vol. 1, no. 1, pp. 233–238, Feb. 1986.

[33] F. Pukelsheim, *Optimal Design of Experiments*. New York: Wiley, 1993.

[34] V. Terzija, G. Valverde, D. Cai, P. Regulski, V. Madani, J. Fitch, S. Skok, M. M. Begovic, and A. Phadke, "Wide-area monitoring, protection, and control of future electric power networks," *Proc. IEEE*, vol. 99, no. 1, pp. 80–93, Jan. 2011.

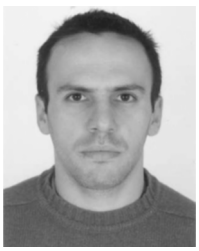
[35] R. H. Tutuncu, K. C. Toh, and M. Todd, "Solving semidefinite-quadratic-linear programs using SDPT3," *Math. Program. Ser. B*, vol. 95, pp. 189–217, 2003.

[36] Power Systems Test Case Archive, Univ. Washington. [Online]. Available: http://www.ee.washington.edu/research/pstca.

[37] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*, 2nd ed. New York: Wiley, 1996.

[38] J. Zhang, G. Welch, and G. Bishop, "Observability and estimation uncertainty analysis for PMU placement alternatives," in *Proc. North American Power Symp. (NAPS)*, Arlington, TX, Sep. 2010.

[39] M. Zhou, V. A. Centeno, J. S. Thorp, and A. G. Phadke, "An alternative for including phasor measurements in state estimators," *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1930–1937, Nov. 2006.
[40] K. Zhu, L. Nordstrom, and L. Ekstam, "Application and analysis of optimum PMU placement methods with application to state estimation accuracy," in *Proc. IEEE PES General Meeting*, Calgary, AB, Canada, Jul. 2009.
[41] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.



Vassilis Kekatos (M'10) was born in Athens, Greece, in 1978. He received the Diploma, M.Sc., and Ph.D. degrees in computer engineering and computer science from the University of Patras, Patras, Greece, in 2001, 2003, and 2007, respectively.

Since 2009, he has been a Marie Curie Fellow, and he is currently a post doctoral associate with the Department of Electrical and Computer Engineering of the University of Minnesota, Minneapolis, and the Computer Engineering and Informatics Department, University of Patras. His research interests lie in the areas of statistical signal processing with emphasis on compressive sampling, wireless communications, and applications for the power grid.

Georgios B. Giannakis (F'97) received the Diploma in electrical engineering from the National Technical University of Athens, Athens, Greece, in 1981, and the M.Sc. degree in electrical engineering, the M.Sc. degree in mathematics, and the Ph.D. degree in electrical engineering from the University of Southern California (USC), Los Angeles, in 1983, 1986, and 1986, respectively.

From 1982 to 1986, he was with USC. Since 1999, he has been a Professor with the University of Minnesota, Minneapolis, where he now holds an ADC Chair in Wireless Telecommunications in the Electrical and Computer Engineering Department, and serves as director of the Digital Technology Center. His general interests span the areas of communications, networking, and statistical signal processing—subjects on which he has published more than 325 journal papers, 525 conference papers, 20 book chapters, two edited books, and two research monographs. Current research focuses on compressive sensing, cognitive radios, network coding, cross-layer designs, wireless sensors, social and power grid networks. He is the (co-) inventor of 21 patents issued.

Dr. Giannakis is the (co-) recipient of eight paper awards from the IEEE Signal Processing (SP) and Communications Societies, including the G. Marconi Prize Paper Award in Wireless Communications. He also received Technical Achievement Awards from the SP Society (2000), from EURASIP (2005), a Young Faculty Teaching Award, and the G. W. Taylor Award for Distinguished Research from the University of Minnesota. He is a Fellow of EURASIP, and has served the IEEE in a number of posts, including that of a Distinguished Lecturer for the IEEE-SP Society.

Bruce Wollenberg (M'67–SM'75–F'89–LF'08) received the B.S.E.E. degree and the M.Eng. degree in electric power engineering from Rensselaer Polytechnic Institute, Troy, NY, in 1964 and 1966, respectively, and the Ph.D. degree in systems engineering from the University of Pennsylvania, Philadelphia, in 1974.

He worked for Leeds and Northrup Co., North Wales, PA from 1966 to 1974, Power Technologies Inc. Schenectady, NY, from 1974 to 1984, and Control Data Corporation Energy Management System Division, Plymouth, MN, from 1984 to 1989. He took a position as a Professor of Electrical Engineering in the Electrical and Computer Engineering Department at the University of Minnesota, Minneapolis, in September 1989. He is presently the Director of the University of Minnesota Center for Electric Energy (UMCEE). His main research interests are the application of mathematical analysis to power system operation and planning problems. He is the co-author of the textbook *Power Generation Operation and Control* (New York: Wiley, 1984).