Online Energy Price Matrix Factorization for Power Grid Topology Tracking

Vassilis Kekatos, *Member, IEEE*, Georgios B. Giannakis, *Fellow, IEEE*, and Ross Baldick, *Fellow, IEEE*

***Abstract*—Grid security and open markets are two major smart grid goals. Transparency of market data facilitates a competitive and efficient energy environment. But it may also reveal critical physical system information. Recovering the grid topology based solely on publicly available market data is explored here. Real-time energy prices are typically calculated as the Lagrange multipliers of network-constrained economic dispatch; that is, via a linear program (LP) typically solved every 5 min. Since the grid Laplacian matrix is a parameter of this LP, someone apart from the system operator could try inferring this topology-related matrix upon observing successive LP dual outcomes. It is first shown that the matrix of spatio-temporal prices can be factored as the product of the inverse Laplacian times a sparse matrix. Leveraging results from sparse matrix decompositions, topology recovery schemes with complementary strengths are subsequently formulated. Solvers scalable to high-dimensional and streaming market data are devised. Numerical validation using synthetic and real-load data on the IEEE 30-bus grid provide useful input for current and future market designs.**

***Index Terms*—Alternating direction method of multipliers (ADMM), compressive sensing, economic dispatch, graph Laplacian, locational marginal prices (LMPs), online convex optimization.**

I. Introduction

AN INDEPENDENT system operator collects energy offers and bids, and dispatches power by maximizing the social welfare while meeting physical grid limitations. To guarantee competitive market operation, multiple data are communicated to market participants or are openly publicized, either in real-time or with certain delay [1]. Such market data may involve energy prices, bids and offers, congestion information, demand, and renewable generation. Looking forward, the smart grid vision calls for energy markets reaching the distribution level to promote participation, accounting for increased stochasticity at a finer time resolution [2]. New reliable market designs are hence to be developed.

From state estimation to load prediction, inference using data has been a major grid operation component. Facing smart grid challenges and opportunities, grid analytics are now extending to price and renewable forecasting, consumer preference learning, and cyber-physical attack detection [3]. Among grid learning tasks, monitoring the transmission network topology is critical for security, market clearing, and billing. Currently, system operators have a precise knowledge of the grid topology using historical data, measurements, and the generalized state estimator [4]. Detecting topological changes from the operator's perspective has been studied in [5] and [6]; while transmission line outages can be efficiently revealed via the sparse overcomplete representation of [7].

Nonetheless, third parties other than the system operator may be interested in tracking transmission network topologies. There has been an increasing concern lately regarding cyber attacks on critical power grid operations. Stealth data attacks were first recognized in [8], while their impact on state estimation and market outcomes was characterized in [9] and [10]. The possibility of data framing attacks deceiving the bad data processor was explored in [11]. Attacks and countermeasures on power system controllers have been studied in [12]. Procedures for detecting and identifying cyber-physical attacks have also been reported (see [13]). In most of these scenarios, designing cyber-physical attacks generally presumes the grid topology to be (locally) known [14], [15]. Apart from data attacks, knowing which transmission lines are (currently) congested could assist in informed bidding or market manipulation [16]. Further, line reactances could be used as a measure of electrical distance to cluster buses, reveal influential nodes, or characterize the performance of decentralized algorithms. Finally, the graph Laplacian of a grid could capture the correlation across pricing nodes [17]. Therefore, recovering this matrix could be capitalized in price forecasting tasks.

Apart from conventional topology tracking via generalized state estimation, alternative schemes have been recently proposed. Grid topology recovery is cast as a blind factorization on the matrix of spatio-temporal power injections in [18]. Even though building on the sparsity and positive semidefiniteness of the grid Laplacian, Li *et al.* [18] relied on linear independence across voltage phases. Considering a power line communication network, time delays of communication signals are leveraged to unveil the microgrid structure in [19]. By postulating a Gaussian Markov random field over bus voltage phases, transmission network faults could also be localized [20]. Likewise for distribution grids, the topology

Manuscript received October 10, 2014; revised March 23, 2015 and June 26, 2015; accepted August 12, 2015. Date of publication August 26, 2015; date of current version April 19, 2016. This work was supported in part by the Institute of Renewable Energy and the Environment under Grant RL-0010-13; in part by the University of Minnesota; and in part by the National Science Foundation under Grant 1423316, Grant 1442686, Grant 1508993, and Grant 1509040. Paper no. TSG-01004-2014.

V. Kekatos is with the Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA 24061 USA (e-mail: kekatos@vt.edu).

G. B. Giannakis is with the Digital Technology Center, Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: georgios@umn.edu).

R. Baldick is with the Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX 78712 USA (e-mail: baldick@ece.utexas.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/TSG.2015.2469098

recovery method of [21] exploits the sample covariance matrix of bus voltage magnitudes.

All in all, existing topology recovery schemes rely on two presumptions: 1) access to a physical system quantity (power injections, voltage phases or magnitudes, communication delays); and 2) this quantity is measured across all buses. Nevertheless, these two assumptions are hard to be satisfied in practice unless the recovery scheme is run by the system operator. In contrast, this paper targets grid topology recovery using readily available market data. The idea here is that real-time energy prices are found by the system operator as the Lagrange multipliers of the network-constrained economic dispatch, which is a linear program (LP) typically solved every 5 min. Dispatch decisions are the primal variables of this LP, while grid topology and electricity offers/bids are its parameters. Observing the dual variables (prices) related to multiple offer/bids instances, the crux is to recover the quasi-stationary topology underlying this LP. Our first contribution is recognizing that properties of the Laplacian matrix and sparsity in congested lines could be exploited: the matrix of spatio-temporal prices can be factorized as the product of a doubly positive matrix with sparse inverse times a sparse matrix. Novel blind recovery schemes of complementary strengths constitute the second contribution: different from [22], the low-rank property of one of the matrix factors is not regularized here, thus significantly simplifying the problem. As our third contribution, algorithms handling big market data are developed based on the alternating direction method of multipliers (ADMM) and its online version. Advancing tools from online optimization, an algorithm handling streaming market data is devised. Distinct from the static approach in [22], this novel online approach is pertinent to future smart grid market designs. Experiments with market data obtained using synthetic prices as well as real-load data over the IEEE 30-bus benchmark corroborate the validity of our findings.

Outline. Section II reviews real-time energy markets and elaborates on the data model. Structural properties and the novel topology recovery are presented in Section III, where two models for energy prices are introduced. Batch solvers for tackling both models are presented in Section IV. Sequential recovery schemes where market data are processed as soon as they are announced are derived in Section V. Section VI shows numerical tests and this paper is concluded in Section VII.

Notation. Lower- (upper-) case boldface letters denote column vectors (matrices); $\mathbf{1}$ and $\mathbf{0}$ denote the all-ones and all-zeros vectors. Symbols $\mathbf{X}'$, $\mathrm{tr}(\mathbf{X})$, and $|\mathbf{X}|$, stand for matrix transposition, trace, and determinant, respectively. Symbol $\mathbb{S}^N$ ($\mathbb{S}^N_+$) is the set of real $N\times N$ symmetric (positive semidefinite) matrices. Regarding matrix norms, $\|\mathbf{A}\|_*$ is the nuclear norm (sum of matrix singular values); $\|\mathbf{A}\|_F$ is the Frobenius norm; and $\|\mathbf{A}\|_1 \coloneqq \sum_{m,n}|\mathbf{A}_{m,n}|$.

II. Energy Price Data Model

Before delineating our price data model, this section reviews linear dc power flows and real-time energy markets.

A. Power Grid Modeling

Consider a power grid represented by the graph $\mathcal{G}=(\mathcal{V},\mathcal{E})$, where the set of nodes $\mathcal{V}$ corresponds to $N+1$ buses, and the edges in $\mathcal{E}$ to L transmission lines. The grid topology is captured via the $L\times(N+1)$ branch-bus incidence matrix $\tilde{\mathbf{A}}$ [3]. For a connected grid, the nullity of $\tilde{\mathbf{A}}$ is one; and by definition, $\tilde{\mathbf{A}}\mathbf{1}=\mathbf{0}$. If $x_l>0$ is the reactance of line l and $\mathbf{D}$ an $L\times L$ diagonal matrix with $[\mathbf{D}]_{l,l}=x_l^{-1}$, the bus reactance matrix can be defined as $\tilde{\mathbf{B}}\coloneqq\tilde{\mathbf{A}}'\mathbf{D}\tilde{\mathbf{A}}$. Given that $\tilde{\mathbf{B}}$ is the weighted Laplacian of $\mathcal{G}$, it is positive semidefinite, and $\mathbf{1}$ is an eigenvector corresponding to $\tilde{\mathbf{B}}$'s zero eigenvalue.

The dc power flow model can now be expressed in matrix-vector form. The active power flow from bus n to bus m over line l can be approximated as $f_l=(\theta_n-\theta_m)/x_l$, where θ_n is the voltage phase at bus n; while the power injection at bus n is $p_n=\sum_{l:(n,m)}f_l$. By stacking $\{\theta_n,p_n\}_{n=1}^{N+1}$ and $\{f_l\}_{l=1}^{L}$ in $\tilde{\boldsymbol{\theta}}$, $\tilde{\mathbf{p}}\in\mathbb{R}^{N+1}$ and $\mathbf{f}\in\mathbb{R}^L$, respectively; it follows that $\mathbf{f}=\mathbf{D}\tilde{\mathbf{A}}\tilde{\boldsymbol{\theta}}$ and $\tilde{\mathbf{p}}=\tilde{\mathbf{A}}'\mathbf{f}=\tilde{\mathbf{B}}\tilde{\boldsymbol{\theta}}$. By eliminating $\tilde{\boldsymbol{\theta}}$, the flows $\mathbf{f}$ can be linearly expressed in terms of $\tilde{\mathbf{p}}$; yet $\tilde{\mathbf{B}}$ is noninvertible.

To resolve the singularity of $\tilde{\mathbf{B}}$, partition $\tilde{\mathbf{A}}$ into the first and the rest of its columns as $\tilde{\mathbf{A}}=[\tilde{\mathbf{a}}\ \mathbf{A}]$. For a connected $\mathcal{G}$, the reduced branch-bus incidence matrix $\mathbf{A}$ has full column-rank. Thus, the reduced bus reactance matrix $\mathbf{B}\coloneqq\mathbf{A}'\mathbf{D}\mathbf{A}$, is strictly positive definite. Setting $\theta_1=0$, it readily follows:

$$\mathbf{f}=\mathbf{T}\tilde{\mathbf{p}} \tag{1}$$

where $\mathbf{T}\coloneqq[\mathbf{0}\ \mathbf{D}\mathbf{A}\mathbf{B}^{-1}]\in\mathbb{R}^{L\times(N+1)}$.

B. Modeling of Real-Time Energy Markets

Building on this model, let us now review real-time energy markets. Energy markets determine the price for electricity by matching supply and demand. Due to time-varying demand and transmission grid limitations, the electricity cost varies across time and space (buses), giving rise to locational marginal prices (LMPs) [1]. Real-time markets are spot markets where hourly power schedules determined over the previous day are adjusted every 5 min to accommodate real-time deviations. Specifically, real-time LMPs are found via the network-constrained economic dispatch, typically formulated as the following LP [23]:

$$\tilde{\mathbf{p}}_t^*\in\arg\min_{\tilde{\mathbf{p}}_t}\quad \tilde{\mathbf{c}}_t'\tilde{\mathbf{p}}_t \tag{2a}$$

$$\text{s.t.}\quad \underline{\mathbf{p}}_t\le\tilde{\mathbf{p}}_t\le\overline{\mathbf{p}}_t \tag{2b}$$

$$\tilde{\mathbf{p}}_t'\mathbf{1}=0 \tag{2c}$$

$$-\bar{\mathbf{f}}\le\mathbf{T}\tilde{\mathbf{p}}_t\le\bar{\mathbf{f}}. \tag{2d}$$

Problem (2) determines the incremental power injections $\tilde{\mathbf{p}}_t^*$ for the upcoming 5-min interval indexed by t. The optimum dispatch $\tilde{\mathbf{p}}_t^*$ is found by minimizing the electricity cost in (2a); while satisfying the power limits in (2b), achieving the supply-demand balance via (2c), and confining line flows approximated by (1) to lie within a secure range [see (2d)]. The power injection bounds in (2b) model bid blocks.

By solving (2), the operator not only determines $\tilde{\mathbf{p}}_t^\star$, but also calculates the LMPs as follows [23]. Let $\lambda_{0,t}$ be the optimal Lagrange multiplier associated with the supply-demand equality in (2c); and $(\underline{\boldsymbol{\mu}}_t,\overline{\boldsymbol{\mu}}_t)\in\mathbb{R}^L_+\times\mathbb{R}^L_+$ be the optimal Lagrange multipliers related to the lower and upper flow limits in (2d). By duality and upon defining $\boldsymbol{\mu}_t\coloneqq\underline{\boldsymbol{\mu}}_t-\overline{\boldsymbol{\mu}}_t$, problem (2) can

be equivalently expressed as

$$\tilde{\mathbf{p}}_t^* \in \arg\max_{\tilde{\mathbf{p}}_t} \quad \left(\lambda_{0,t}\mathbf{1}+\mathbf{T}'\boldsymbol{\mu}_t-\tilde{\mathbf{c}}_t\right)'\tilde{\mathbf{p}}_t \quad \text{s.t.} \quad \underline{\mathbf{p}}_t \le \tilde{\mathbf{p}}_t \le \overline{\mathbf{p}}_t. \tag{3}$$

If $\lambda_{0,t}\mathbf{1}+\mathbf{T}'\boldsymbol{\mu}_t$ is selected as the vector of bus electricity prices at time t and assuming $\tilde{\mathbf{c}}_t$ are the actual marginal costs, then (3) reveals that $\tilde{\mathbf{p}}_t^\star$ maximizes the sum of individual profits. Note that if a different bus is selected as reference, the corresponding Lagrange multipliers will be different from $(\lambda_{0,t}, \underline{\boldsymbol{\mu}}_t, \overline{\boldsymbol{\mu}}_t)$. However, the value of the equivalent vector to $\lambda_{0,t}\mathbf{1}+\mathbf{T}'\boldsymbol{\mu}_t$ will be the same; thus indicating that the price vector $\lambda_{0,t}\mathbf{1}+\mathbf{T}'\boldsymbol{\mu}_t$ is invariant to the reference bus choice.

Recall that the approximate dc model ignores the heat losses dissipated on transmission lines. Once the operator has found $\lambda_{0,t}\mathbf{1}+\mathbf{T}'\boldsymbol{\mu}_t$ via (2), it modifies prices to account for the cost of power losses. Thus, LMPs are finally calculated as

$$\tilde{\boldsymbol{\pi}}_t := \lambda_{0,t}\mathbf{1}+\begin{bmatrix}0\\ \mathbf{B}^{-1}\mathbf{A}'\mathbf{D}\boldsymbol{\mu}_t\end{bmatrix}+\tilde{\boldsymbol{\ell}}_t \tag{4}$$

where $\tilde{\boldsymbol{\ell}}_t$ is a loss correction [23]. The LMPs in (4) consist of three summands: 1) the marginal energy component (MEC) $\lambda_{0,t}$; 2) the marginal congestion component (MCC) $[0\ \ \boldsymbol{\mu}_t'\mathbf{D}\mathbf{A}\mathbf{B}^{-1}]'$; and 3) the marginal loss component (MLC) $\tilde{\boldsymbol{\ell}}_t$. Starting from the last, the MLC is currently calculated in different ways by different system operators. Nonetheless, given that heat losses in transmission grids are typically 3%–5% of energy consumption, the component $\tilde{\boldsymbol{\ell}}_t$ is relatively small compared to the other two. For example, for the mid-continent independent system operator market in March 11, 2015 at 6 P.M. the network-averaged absolute LMP was 24\$/MWh, whereas the network-averaged absolute MLC was only 0.7\$/MWh.

According to (3), the MEC is the energy price at the reference bus. When the power flow on line l reaches the upper or lower limit at time t, that is $f_{l,t}=\overline{f}_l$ or $f_{l,t}=-\overline{f}_l$, then line l is termed congested. The ensuing fact for the $\boldsymbol{\mu}_t$'s follows trivially from complementary slackness.

Proposition 1: If transmission line l is not congested at time t, then the lth entry of $\boldsymbol{\mu}_t$ is zero.

Apparently, if there are no congested lines and losses were ignored, all nodes would enjoy the same energy price $\lambda_{0,t}$.

C. Problem Statement

Depending on the market, the LMP components are announced either separately or collectively as a sum. In the former case, the MCC is readily available. If the sum of MEC and MCC is publicized, the effect of MEC can be isolated by subtracting the price of the first entry of $\tilde{\boldsymbol{\pi}}_t$ from all entries of $\tilde{\boldsymbol{\pi}}_t$. It can be argued that subtracting the first entry does not harm the generality of this preprocessing step regarding the sum of MEC and MCC, even if the reference bus is not bus 1, since the assumed model $\lambda_{0,t}\mathbf{1}+\mathbf{T}'\boldsymbol{\mu}_t$ still holds. Either way, collect all but the first bus prices in $\boldsymbol{\pi}_t \in \mathbb{R}^N$, for which we postulate

$$\boldsymbol{\pi}_t = \mathbf{B}^{-1}\mathbf{s}_t+\mathbf{n}_t \tag{5}$$

where $\mathbf{s}_t := \mathbf{A}'\mathbf{D}\boldsymbol{\mu}_t$ and $\mathbf{n}_t$ captures the MLC (in its original or its preprocessed form). Slightly abusing terminology, $\boldsymbol{\pi}_t$ will be henceforth termed the LMPs.

Market clearing occurs every 5 min, and market bids $\{\tilde{\mathbf{c}}_t, \underline{\mathbf{p}}_t, \overline{\mathbf{p}}_t\}$ change partially over time, adapting to demand and generation fluctuations. Consider collecting the LMPs of (5) over the horizon $\mathcal{T} := \{t : t = 1, \ldots, T\}$ of T consecutive market intervals, and suppose grid topology remains invariant over $\mathcal{T}$. Upon stacking $\{\boldsymbol{\pi}_t, \mathbf{s}_t, \mathbf{n}_t\}_{t\in\mathcal{T}}$ as columns of the $N\times T$ matrices $\boldsymbol{\Pi}$, $\mathbf{S}$, and $\mathbf{N}$, respectively, it follows from (5):

$$\boldsymbol{\Pi} = \mathbf{B}^{-1}\mathbf{S}+\mathbf{N}. \tag{6}$$

Model (6) asserts that if $\mathbf{N}$ is ignored, the price matrix $\boldsymbol{\Pi}$ can be factorized as the product of the inverse grid Laplacian $\mathbf{B}^{-1}$ times matrix $\mathbf{S}$. With (6), topology recovery can be now formulated as the problem of finding $(\mathbf{B}, \mathbf{S})$ given $\{\boldsymbol{\pi}_t\}_{t\in\mathcal{T}}$.

Remark 1: Having multiblock offers and several bidders per bus does not harm generality of (4)–(6). Specifically, electricity offers and bids oftentimes consist of multiple blocks. For example, a generator may offer to produce the first 20 MWh for at least 20\$/MWh and the next 5 MWh for at least 23\$/MWh at the same bus. In this case, the corresponding $p_{n,t}$ should be decomposed as the sum of two extra optimization variables as $p_{n,t} = p_{n,t}^1 + p_{n,t}^2$ with $0 \le p_{n,t}^1 \le 20$ and $20 \le p_{n,t}^2 \le 25$; while the summand $c_{n,t}p_{n,t}$ in (2) is replaced by $20p_{n,t}^1 + 23p_{n,t}^2$. Having multiple generators and/or consumers at the same bus is handled similarly. Either way, constraints (2c)–(2d) apparently remain unaltered. Hence, even though simplifying, (2) is sufficiently representative. Actually, Section VI involves tests with multiblock offers.

III. Topology Recovery Approaches

A. Noiseless Price Model

If the MCCs are announced separately from the MLCs, matrix $\boldsymbol{\Pi}$ satisfies the noiseless counterpart of model (6)

$$\boldsymbol{\Pi} = \mathbf{B}^{-1}\mathbf{S}. \tag{7}$$

Decomposing $\boldsymbol{\Pi}$ into $(\mathbf{B}, \mathbf{S})$ constitutes a blind matrix factorization problem. To uniquely recover the factors, their rich structure delineated next should be properly exploited.

Recall that $\mathbf{B}$ is positive definite. Once $\mathbf{B}$ has been recovered, the original grid Laplacian $\tilde{\mathbf{B}}$ can be trivially found in light of the property $\tilde{\mathbf{B}}\mathbf{1} = \mathbf{0}$. Note further that the (n, m)-th entry of $\mathbf{B}$ equals $-x_{nm}^{-1}$, if there is a line between buses m and n; and zero otherwise. Granted power grids are sparingly connected, $\mathbf{B}$ is not only sparse, but its off-diagonal entries are nonpositive. Having positive eigenvalues and nonpositive off-diagonal entries implies $\mathbf{B}$ is an invertible M-matrix [24, Sec. 2.5]. Hence, $\mathbf{B}^{-1}$ has positive entries, i.e., $\mathbf{B}^{-1} > \mathbf{0}$.

As far as $\mathbf{S}$ is concerned, its columns can be expressed as $\mathbf{s}_t = \sum_{l\in\mathcal{E}} \mu_{t,l}x_l^{-1}\mathbf{a}_l$. Since many of the $\{\mu_{t,l}\}_l$ in (8) are expected to be zero (see Proposition 1), $\mathbf{s}_t$ can also be written as

$$\mathbf{s}_t = \sum_{l\in\mathcal{C}_t} \frac{\mu_{t,l}}{x_l}\mathbf{a}_l \tag{8}$$

where $\mathcal{C}_t \subseteq \mathcal{E}$ is the subset of congested lines at time t. In other words, $\mathbf{s}_t$ is a linear combination of few $\mathbf{a}_l$'s. Given that $\mathbf{a}_l$ are sparse, matrix $\mathbf{S}$ is expected to be sparse too. Typically, only a

few transmission lines become congested over a short market period (say one day). In the California ISO, for example, only two transmission lines are typically congested [25]. Hence, it can be assumed that the $\{\mathcal{C}_t\}_{t=1}^T$ overlap significantly, or that the locations of the nonzero entries of $\{\boldsymbol{\mu}_t\}_{t=1}^T$ remain relatively time-invariant. Since $\mathbf{s}_t$'s are linear combinations of those few $\mathbf{a}_l$'s corresponding to congested lines, $\mathbf{S}$ is expected to exhibit low rank. The invertibility of $\mathbf{B}$ implies $\mathbf{\Pi}$ should have low rank too.

It will also be useful to recognize that the factorization in (7) is scale-invariant: if $(\mathbf{B}, \mathbf{S})$ satisfies (7), so does the pair $(\alpha\mathbf{B}, \alpha\mathbf{S})$ for all $\alpha > 0$. To waive this inherent ambiguity, the maximum diagonal entry of $\mathbf{B}$ is assumed to be unity. Due to this normalization, matrix $\mathbf{B}$ should satisfy $\mathbf{B} \succeq \mathbf{0}$ and $\mathbf{B} \leq \mathbf{I}$.

Leveraging these properties, one could recover $(\mathbf{B}, \mathbf{S})$ by solving the optimization problem:

$$\min_{\mathbf{B},\mathbf{S}} \quad \|\mathbf{S}\|_0 + \kappa_0\|\mathbf{B}\|_0 \tag{9a}$$

$$\text{s.t.} \quad \mathbf{B\Pi} = \mathbf{S},\ \mathbf{B} \succ \mathbf{0},\ \mathbf{B} \leq \mathbf{I} \tag{9b}$$

where $\|\mathbf{X}\|_0$ is the ℓ_0-(pseudo)norm of a matrix counting its nonzero entries, and $\kappa_0 > 0$ is a parameter balancing the sparsity between the two matrices. Problem (9) finds the sparsest pair $(\mathbf{B}, \mathbf{S})$ that satisfies model (7) and the structural constraints of $\mathbf{B}$. Different from [22], the rank of $\mathbf{S}$ is not penalized here, since the constraint $\mathbf{B\Pi} = \mathbf{S}$ and the nonsingularity of $\mathbf{B}$ imply $\text{rank}(\mathbf{S}) = \text{rank}(\mathbf{\Pi})$ anyway.

Minimizing the ℓ_0-norm is in general NP-hard [26]. Following advances in compressive sensing [27], the ℓ_0-norm will be surrogated by the ℓ_1-norm to yield the convex problem:

$$\min_{\mathbf{B},\mathbf{S}} \quad \|\mathbf{S}\|_1 + \kappa_1\,\text{tr}(\mathbf{PB}) - \kappa_2 \log|\mathbf{B}| \tag{10}$$

$$\text{s.t.} \quad \mathbf{B\Pi} = \mathbf{S},\ \mathbf{B} \in \mathcal{B}$$

where $\mathbf{P} := \mathbf{I} - \mathbf{1}\mathbf{1}'$, $\mathcal{B} := \{\mathbf{B} : \mathbf{B} \succeq \mathbf{0},\ \mathbf{B} \leq \mathbf{I}\}$, and $\kappa_1, \kappa_2 > 0$. Two observations are in order regarding (10).

First, since that the diagonal entries of $\mathbf{B}$ are strictly positive, $\|\mathbf{B}\|_0$ in (9a) has been replaced by the sum of the off-diagonal entries of $\mathbf{B}$ in their absolute values, that is

$$\sum_{n,m:n\neq m} |\mathbf{B}_{n,m}| = \sum_n \mathbf{B}_{n,n} - \sum_{n,m} \mathbf{B}_{n,m} = \text{tr}(\mathbf{B}) - \mathbf{1}'\mathbf{B}\mathbf{1} = \text{tr}(\mathbf{PB})$$

where the first equality comes from the nonpositive off-diagonal entries of $\mathbf{B}$, and the rest from properties of the trace.

Second, ideally $\mathbf{B}$ should be enforced to be strictly positive definite, i.e., $\mathbf{B} \succ \mathbf{0}$. Nonetheless, current optimization algorithms cannot guarantee the minimizer to lie in the interior of the feasible set. On the other hand, imposing $\mathbf{B} \succeq \mathbf{0}$ admits the trivial solution $(\mathbf{B}, \mathbf{S}) = (\mathbf{0}, \mathbf{0})$. As a remedy, the term $-\log|\mathbf{B}|$ has been added in the cost of (10) to confine $\mathbf{B}$ in the interior of the positive semidefinite cone $\mathbf{B} \succeq \mathbf{0}$.

By eliminating $\mathbf{S}$, (10) can be equivalently transformed to

$$\min_{\mathbf{B}\in\mathcal{B}} \quad \|\mathbf{B\Pi}\|_1 + \kappa_1\,\text{tr}(\mathbf{PB}) - \kappa_2 \log|\mathbf{B}|. \tag{11}$$

The strict convexity of $-\kappa_2 \log|\mathbf{B}|$ guarantees that (11) and hence (10) have unique minimizers.

B. Noisy Price Model

In practice, the price model in (6) may be more pertinent than the exact model of (7). That is the case for example when $\{\boldsymbol{\pi}_t\}_{t\in\mathcal{T}}$ comprise of both congestion (MCC) and loss (MLC) components. Furthermore, model (6) captures perturbations in $\mathbf{B}$, which could be attributed to the linearization of the original ac power flow model or to time-varying transformer settings. Either way, the nonzero entries of $\mathbf{B}$ are perturbed according to the instantaneous grid operating point.

Rather than enforcing the constraint $\mathbf{B\Pi} = \mathbf{S}$, the idea here is to look for sparse and nonzero $(\mathbf{B}, \mathbf{S})$ that yield a small least-squares error $\|\mathbf{B\Pi} - \mathbf{S}\|_F^2$. Hence, the nonconvex problem in (9) is replaced by

$$\min_{\mathbf{B},\mathbf{S}} \quad \tfrac{1}{2}\|\mathbf{B\Pi} - \mathbf{S}\|_F^2 + \kappa_0\|\mathbf{B}\|_0 + \kappa_0'\|\mathbf{S}\|_0 \tag{12}$$

$$\text{s.t.} \quad \mathbf{B} \succ \mathbf{0},\ \mathbf{B} \leq \mathbf{I}$$

for $\kappa_0, \kappa_0' > 0$. In the same spirit (9) was surrogated by (10), the hard problem in (12) is approximated by the convex one

$$\min_{\mathbf{B}\in\mathcal{B},\mathbf{S}} \tfrac{1}{2}\|\mathbf{B\Pi} - \mathbf{S}\|_F^2 + \kappa_3\|\mathbf{S}\|_1 + \kappa_1\,\text{tr}(\mathbf{PB}) - \kappa_2 \log|\mathbf{B}| \tag{13}$$

for some $\kappa_3 > 0$. Actually, after minimizing over $\mathbf{S}$, the last problem can be shown to be equivalent to

$$\min_{\mathbf{B}\in\mathcal{B}} \quad \tilde{h}_{\kappa_3}(\mathbf{B\Pi}) + \kappa_1\,\text{tr}(\mathbf{PB}) - \kappa_2 \log|\mathbf{B}| \tag{14}$$

where $\tilde{h}_{\kappa_3}(\mathbf{X}) := \sum_{m,n} h_{\kappa_3}(\mathbf{X}_{m,n})$, and h_{κ_3} is the so termed Huber function

$$h_{\kappa_3}(x) := \begin{cases} \dfrac{1}{2}x^2, & |x| \leq \kappa_3 \\ \kappa_3|x| - \dfrac{\kappa_3^2}{2}, & |x| > \kappa_3. \end{cases} \tag{15}$$

Comparing (11) with (14), the entries of $\mathbf{B\Pi}$ in (11) are arguments of the absolute value cost. On the other hand, the objective in (14) penalizes small entries of $\mathbf{B\Pi}$ with a quadratic cost, and large ones with the absolute value cost [see (15)].

IV. Batch Topology Recovery Schemes

Although (10) and (13) could be solved by commercial software for relatively small problems, interior point-based solvers cannot handle N and T larger than a few hundreds. There are two main optimization challenges here: 1) the nondifferentiable objective terms $\|\mathbf{B\Pi}\|_1$ or $\|\mathbf{S}\|_1$; and 2) the feasible set $\mathcal{B}$. Note that $\mathcal{B}$ is the intersection of the positive definite cone and a shifted version of the nonnegative orthant. Albeit projection over each of these convex cones is relatively easy, there is no closed-form solution for projecting on $\mathcal{B}$.

Given these challenges, an algorithm based on the ADMM is derived next. ADMM targets solving problems of the form [28]

$$\min_{\mathbf{x}\in\mathcal{X},\mathbf{z}\in\mathcal{Z}} \quad \{f(\mathbf{x}) + g(\mathbf{z}) : \mathbf{F}\mathbf{x} + \mathbf{G}\mathbf{z} = \mathbf{c}\} \tag{16}$$

where $f(\mathbf{x})$ and $g(\mathbf{z})$ are convex functions; $\mathcal{X}$ and $\mathcal{Z}$ are convex sets; and $(\mathbf{F}, \mathbf{G}, \mathbf{c})$ model the linear equality constraints

coupling variables $\mathbf{x}$ and $\mathbf{z}$. In its normalized form, ADMM assigns a Lagrange multiplier $\mathbf{y}$ for the equality constraint and solves (16) by iterating over the recursions

$$\mathbf{x}^{i+1} := \arg\min_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x}) + \tfrac{\rho}{2}\left\|\mathbf{F}\mathbf{x}+\mathbf{G}\mathbf{z}^{i}-\mathbf{c}+\mathbf{y}^{i}\right\|_2^2 \tag{17a}$$

$$\mathbf{z}^{i+1} := \arg\min_{\mathbf{z}\in\mathcal{Z}} g(\mathbf{z}) + \tfrac{\rho}{2}\left\|\mathbf{F}\mathbf{x}^{i+1}+\mathbf{G}\mathbf{z}-\mathbf{c}+\mathbf{y}^{i}\right\|_2^2 \tag{17b}$$

$$\mathbf{y}^{i+1} := \mathbf{y}^{i} + \mathbf{F}\mathbf{x}^{i+1} + \mathbf{G}\mathbf{z}^{i+1} - \mathbf{c} \tag{17c}$$

for some $\rho > 0$. The generic algorithmic steps of (17) are engaged in solving (10) and (13) next.

A. Batch Solver for the Noiseless Model

To apply ADMM and end up in efficient updates for (10), we first replace variable $\mathbf{B}$ with three copies $\mathbf{B}_1$, $\mathbf{B}_2$, and $\mathbf{B}_3$, to yield the equivalent problem

$$\min_{\mathbf{B}_1,\mathbf{B}_2\le\mathbf{I},\mathbf{B}_3\ge\mathbf{0},\mathbf{S}} \quad \|\mathbf{S}\|_1 + \kappa_1\operatorname{tr}(\mathbf{P}\mathbf{B}_1) - \kappa_2\log|\mathbf{B}_3| \tag{18a}$$

$$\text{s.t.} \quad \mathbf{B}_1 = \mathbf{B}_2 \tag{18b}$$

$$\mathbf{B}_1 = \mathbf{B}_3 \tag{18c}$$

$$\mathbf{B}_1\mathbf{\Pi} = \mathbf{S}. \tag{18d}$$

Let $\mathbf{M}_{12}$, $\mathbf{M}_{13}$, and $\mathbf{M}$, denote the Lagrange multipliers corresponding to (18b), (18c), and (18d), respectively. Partitioning variables into $\mathbf{B}_1$ and $(\mathbf{B}_2, \mathbf{B}_3, \mathbf{S})$, ADMM iterates through the next three steps.

At the first step of iterate i, the variable $\mathbf{B}_1$ is updated given $(\mathbf{B}_2^i, \mathbf{B}_3^i, \mathbf{S}^i)$ and $(\mathbf{M}_{12}^i, \mathbf{M}_{13}^i, \mathbf{M}^i)$ by solving

$$\min_{\mathbf{B}_1}\ \Big\{\kappa_1\operatorname{tr}(\mathbf{P}\mathbf{B}_1) + \tfrac{\rho}{2}\left\|\mathbf{B}_1-\mathbf{B}_2^i+\mathbf{M}_{12}^i\right\|_F^2 + \tfrac{\rho}{2}\left\|\mathbf{B}_1-\mathbf{B}_3^i+\mathbf{M}_{13}^i\right\|_F^2 + \tfrac{\rho}{2}\left\|\mathbf{B}_1\mathbf{\Pi}-\mathbf{S}^i+\mathbf{M}^i\right\|_F^2\Big\}. \tag{19}$$

The solution of (19) is provided in closed form as $\mathbf{B}_1^{i+1} = (\mathbf{B}_2^i-\mathbf{M}_{12}^i+\mathbf{B}_3^i-\mathbf{M}_{13}^i+(\mathbf{S}^i-\mathbf{M}^i)\mathbf{\Pi}'-(\kappa_1/\rho)\mathbf{P})\left(2\mathbf{I}+\mathbf{\Pi}\mathbf{\Pi}'\right)^{-1}$.

During the second step, ADMM updates the second block of variables $(\mathbf{B}_2, \mathbf{B}_3, \mathbf{S})$ given $\mathbf{B}_1^{i+1}$ and $(\mathbf{M}_{12}^i, \mathbf{M}_{13}^i, \mathbf{M}^i)$. Yet the optimization decouples over the three variables. Specifically, variable $\mathbf{B}_2$ is updated as the solution of

$$\min_{\mathbf{B}_2\le\mathbf{I}} \quad \tfrac{\rho}{2}\left\|\mathbf{B}_1^{i+1}-\mathbf{B}_2+\mathbf{M}_{12}^i\right\|_F^2 \tag{20}$$

whose minimizer is $\mathbf{B}_2^{i+1} = \min\{\mathbf{B}_1^{i+1}+\mathbf{M}_{12}^i, \mathbf{I}\}$, where the minimum operator is understood entry-wise.

Variable $\mathbf{B}_3$ is updated as the minimizer of

$$\min_{\mathbf{B}_3\ge\mathbf{0}} \quad \tfrac{1}{2}\left\|\mathbf{B}_1^{i+1}-\mathbf{B}_3+\mathbf{M}_{13}^i\right\|_F^2 - \tfrac{\kappa_2}{\rho}\log|\mathbf{B}_3| \tag{21}$$

which can be readily found as follows [22, Lemma 1]: define the operator $\mathcal{P}_\alpha : \mathbb{R}^{N\times N} \to \mathbb{S}_+^N$ for some $\alpha > 0$ as:

$$\mathcal{P}_\alpha(\mathbf{X}) = \tfrac{1}{2}\mathbf{V}\left(\mathbf{\Xi} + \left(\mathbf{\Xi}^2+4\alpha\mathbf{I}\right)^{1/2}\right)\mathbf{V}' \tag{22}$$

where $\mathbf{V}\mathbf{\Xi}\mathbf{V}'$ is the eigenvalue decomposition of the symmetric matrix $(1/2)(\mathbf{X}+\mathbf{X}')$. Then, the solution to (21) is

$$\mathbf{B}_3^{i+1} = \mathcal{P}_{\kappa_2/\rho}\left(\mathbf{B}_1^{i+1}+\mathbf{M}_{13}^i\right). \tag{23}$$

Algorithm 1 Batch Topology Recovery (Noiseless Model)

Require: Price matrix $\mathbf{\Pi}$ and $(\kappa_1, \kappa_2, \rho)$.
1: Initialize $\mathbf{B}_1^0 = \mathbf{B}_2^0 = \mathbf{B}_3^0 = \mathbf{I}_N$ and $\mathbf{S}^0 = \mathbf{\Pi}$.
2: Initialize $\mathbf{M}_{12}^0 = \mathbf{M}_{13}^0 = \mathbf{0}_N$ and $\mathbf{M}^0 = \mathbf{0}_{N\times T}$.
3: **for** $i = 1, 2, \ldots,$ **do**
4: Update $\mathbf{B}_1^{i+1}$ from (19).
5: Update $\mathbf{B}_2^{i+1}$ from (20).
6: Update $\mathbf{B}_3^{i+1}$ from (23).
7: Update $\mathbf{S}^{i+1}$ from (25).
8: Update multipliers $(\mathbf{M}_{12}^{i+1}, \mathbf{M}_{13}^{i+1}, \mathbf{M}^{i+1})$ from (26).
9: **end for**

Variable $\mathbf{S}$ is updated by solving

$$\min_{\mathbf{S}} \quad \tfrac{1}{2}\left\|\mathbf{B}_1^{i+1}\mathbf{\Pi}-\mathbf{S}+\mathbf{M}^i\right\|_F^2 + \tfrac{1}{\rho}\|\mathbf{S}\|_1. \tag{24}$$

Problem (24) is separable across the entries of $\mathbf{S}$, admitting the closed-form minimizer

$$\mathbf{S}^{i+1} = \mathcal{S}_{\rho^{-1}}\left(\mathbf{B}_1^{i+1}\mathbf{\Pi}+\mathbf{M}^i\right) \tag{25}$$

where $\mathcal{S}_\beta(x)$ is the soft thresholding operator defined as

$$\mathcal{S}_\beta(x) := \begin{cases} x-\beta\operatorname{sign}(x), & |x|>\beta \\ 0, & |x|\le\beta \end{cases}$$

applied entry-wise in (25).

In the third step, the Lagrange multipliers are updated as

$$\begin{aligned}\mathbf{M}_{12}^{i+1} &= \mathbf{M}_{12}^i + \mathbf{B}_1^{i+1} - \mathbf{B}_2^{i+1}\\ \mathbf{M}_{13}^{i+1} &= \mathbf{M}_{13}^i + \mathbf{B}_1^{i+1} - \mathbf{B}_3^{i+1}\\ \mathbf{M}^{i+1} &= \mathbf{M}^i + \mathbf{B}_1^{i+1}\mathbf{\Pi} - \mathbf{S}^{i+1}.\end{aligned} \tag{26}$$

The algorithm is summarized as Algorithm 1, and its convergence is guaranteed for all $\rho > 0$ [28]. Its computational complexity per ADMM iteration is $\mathcal{O}(N^3 + N^2T)$.

B. Batch Solver for the Noisy Model

Finding an efficient ADMM solver for (13) proceeds in a similar manner. To decouple the constraints and the nondifferentiable summands in the cost, introduce three copies for variable $\mathbf{B}$ and two for $\mathbf{S}$ to transform (13) into

$$\min_{\mathbf{B}_1,\mathbf{B}_2\le\mathbf{I},\mathbf{B}_3\ge\mathbf{0},\mathbf{S}_1,\mathbf{S}_2} \quad \tfrac{1}{2}\|\mathbf{B}_1\mathbf{\Pi}-\mathbf{S}_2\|_F^2 + \kappa_3\|\mathbf{S}_1\|_1 + \kappa_1\operatorname{tr}(\mathbf{P}\mathbf{B}_1) - \kappa_2\log|\mathbf{B}_3| \tag{27a}$$

$$\text{s.t.} \quad \mathbf{B}_1 = \mathbf{B}_2 \tag{27b}$$

$$\mathbf{B}_1 = \mathbf{B}_3 \tag{27c}$$

$$\mathbf{S}_1 = \mathbf{S}_2. \tag{27d}$$

Define the Lagrange multipliers $\mathbf{M}_{12}$, $\mathbf{M}_{13}$, and $\mathbf{N}_{12}$, related to equality constraints (27b), (27c), and (27d), respectively; and partition variables as $(\mathbf{B}_1, \mathbf{S}_1)$ and $(\mathbf{B}_2, \mathbf{B}_3, \mathbf{S}_2)$. Following similar steps as in Section IV-A, the ADMM iterations listed in Algorithm 2 can be derived. Its computational complexity per ADMM iteration is $\mathcal{O}(N^3 + N^2T)$.

Algorithm 2 Batch Topology Recovery (Noisy Model)

Require: Price matrix $\mathbf{\Pi}$ and $(\kappa_1, \kappa_2, \kappa_3, \rho)$.

Initialize $\mathbf{B}_1^0 = \mathbf{B}_2^0 = \mathbf{B}_3^0 = \mathbf{I}_N$ and $\mathbf{S}_1^0 = \mathbf{S}_2^0 = \mathbf{\Pi}$.

Initialize $\mathbf{M}_{12}^0 = \mathbf{M}_{13}^0 = \mathbf{0}_N$ and $\mathbf{N}_{12}^0 = \mathbf{0}_{N\times T}$.

for $i = 1, 2, \ldots,$ **do**

$\mathbf{B}_1^{i+1} := (\mathbf{S}_2^i\mathbf{\Pi}' + \rho(\mathbf{B}_2^i - \mathbf{M}_{12}^i) - \kappa_1\mathbf{P})\big(\mathbf{\Pi}\mathbf{\Pi}' + \rho\mathbf{I}_N\big)^{-1}$.

$\mathbf{S}_1^{i+1} := \mathcal{S}_{\kappa_3/\rho}\big(\mathbf{S}_2^i - \mathbf{N}_{12}^i\big)$.

Update $\mathbf{B}_2^{i+1}$ from (20).

Update $\mathbf{B}_3^{i+1}$ from (23).

$\mathbf{S}_2^{i+1} := \frac{1}{\rho+1}\Big(\mathbf{B}_1^{i+1}\mathbf{\Pi} + \rho(\mathbf{S}_1^{i+1} + \mathbf{N}_{12}^i)\Big)$

$\mathbf{M}_{12}^{i+1} := \mathbf{B}_1^{i+1} - \mathbf{B}_2^{i+1}$

$\mathbf{M}_{13}^{i+1} := \mathbf{B}_1^{i+1} - \mathbf{B}_3^{i+1}$

$\mathbf{N}_{12}^{i+1} := \mathbf{S}_1^{i+1} - \mathbf{S}_2^{i+1}$

end for

V. Grid Topology Tracking

The topology recovery scheme of Section IV presumes that: (c1) the power system topology remains unchanged and (c2) prices are available over the entire period $\mathcal{T}$. Yet future power grids may be reconfigured frequently for dispatching and maintenance, while real-time LMPs over thousands of buses are expected to be announced at a fast rate, which render conditions (c1) and (c2) unrealistic. Even for contemporary energy markets, one would like to update an existing grid topology estimate every time a new LMP vector $\boldsymbol{\pi}_t$ arrives. Such a sequential scheme not only waives a multiple-period delay T, but it further relaxes the stationary topology assumption, thus enabling topology tracking.

Solvers for streaming rather than batch pricing data are developed next. The desiderata here is online schemes where topology estimates $\mathbf{B}^t$ are updated every time a price vector $\boldsymbol{\pi}_t$ is publicized. Advances from online optimization are particularly suitable for this task [29]. Tailored to big data processing, many online convex optimization algorithms aim at solving problems of the form

$$\min_{\mathbf{x}\in\mathcal{X}} \quad \sum_{t=1}^{T}(f_t(\mathbf{x}) + g(\mathbf{x})) \tag{28}$$

where $f_t(\mathbf{x})$ depends on the tth datum, and $g(\mathbf{x})$ is a regularizer, i.e., a function leveraging prior information on $\mathbf{x}$.

Tailoring our topology recovery task to the online optimization setup, consider the general problem

$$\min_{\mathbf{B}\in\mathcal{B}} \quad \sum_{t=1}^{T}\big(f_{\boldsymbol{\pi}_t}(\mathbf{B}) + \tfrac{\kappa_1}{T}\operatorname{tr}(\mathbf{PB}) - \tfrac{\kappa_2}{T}\log|\mathbf{B}|\big). \tag{29}$$

By selecting $f_{\boldsymbol{\pi}_t}(\mathbf{B}) := \|\mathbf{B}\boldsymbol{\pi}_t\|_1$, problem (29) yields (11); whereas, for $f_{\boldsymbol{\pi}_t}(\mathbf{B}) := \tilde{h}_{\kappa_3}(\mathbf{B}\boldsymbol{\pi}_t)$, (29) is equivalent to (14). Apparently, the term $f_{\boldsymbol{\pi}_t}(\mathbf{B})$ depends on market data $\boldsymbol{\pi}_t$ and can be thought of as a data fittng term. The other two terms in the objective of (29) can be interpreted as regularizers for $\mathbf{B}$.

To solve (29) in an online fashion, we resorted to the online ADMM algorithm of [28] that cycles through

$$\mathbf{x}^{t+1} := \arg\min_{\mathbf{x}\in\mathcal{X}} f_t(\mathbf{x}) + \tfrac{\rho}{2}\big\|\mathbf{F}\mathbf{x} + \mathbf{G}\mathbf{z}^t - \mathbf{c} + \mathbf{y}^t\big\|_2^2 + \tfrac{\eta}{2}\big\|\mathbf{x} - \mathbf{x}^t\big\|_2^2 \tag{30a}$$

$$\mathbf{z}^{t+1} := \arg\min_{\mathbf{z}\in\mathcal{Z}} g(\mathbf{z}) + \tfrac{\rho}{2}\Big\|\mathbf{F}\mathbf{x}^{t+1} + \mathbf{G}\mathbf{z} - \mathbf{c} + \mathbf{y}^t\Big\|_2^2 \tag{30b}$$

$$\mathbf{y}^{t+1} := \mathbf{y}^t + \Big(\mathbf{F}\mathbf{x}^{t+1} + \mathbf{G}\mathbf{z}^{t+1} - \mathbf{c}\Big). \tag{30c}$$

Comparing (30) with its batch counterpart in (17), the iteration index i in (30) coincides with the data index t, while the first step in (30a) entails only the current $f_t(\mathbf{x})$ together with the proximal term $(\eta/2)\|\mathbf{x} - \mathbf{x}^t\|_2^2$ for some $\eta > 0$. Regarding its convergence, the algorithm attains sublinear regret in terms of both the cost and the constraint violation [28, Th. 4].

Building on (18), introduce copies of $\mathbf{B}$ to express (29) as

$$\min_{\mathbf{B}_1,\mathbf{B}_2,\mathbf{B}_3} \quad \sum_{t=1}^{T}\big(f_{\boldsymbol{\pi}_t}(\mathbf{B}_1) + \tfrac{\kappa_1}{T}\operatorname{tr}(\mathbf{PB}_1) - \tfrac{\kappa_2}{T}\log|\mathbf{B}_3|\big) \tag{31a}$$

$$\text{s.t.} \quad \mathbf{B}_1 = \mathbf{B}_2 \tag{31b}$$

$$\mathbf{B}_1 = \mathbf{B}_3 \tag{31c}$$

$$\mathbf{B}_3 \succeq \mathbf{0}, \ \ \mathbf{B}_2 \preceq \mathbf{I}. \tag{31d}$$

It is worth stressing that even though the batch ADMM solvers of Section IV addressed (10) and (13), our approach in (31) reformulates the problem versions (11) and (14) where variable $\mathbf{S}$ has been eliminated. This is possible only because the online version of ADMM treats each $f_{\boldsymbol{\pi}_t}(\mathbf{B})$ separately. This critical feature enables neat closed-form updates for $\mathbf{B}$ without having to deal with $\mathbf{S}$ as detailed next.

Similarly to (18), let $\mathbf{M}_{12}$ and $\mathbf{M}_{13}$ be the Lagrange multipliers corresponding to constraints (31b) and (31c), respectively. As soon as $\boldsymbol{\pi}_t$ is announced, a cycle of the online ADMM of (30) is initiated. In its first step, $\mathbf{B}_1$ is updated via (30a), which upon completing the squares yields

$$\mathbf{B}_1^t := \arg\min_{\mathbf{B}_1} \ \tfrac{1}{2\rho+\eta}f_{\boldsymbol{\pi}_t}(\mathbf{B}_1) + \tfrac{1}{2}\Big\|\mathbf{B}_1 - \check{\mathbf{B}}_1^{t-1}\Big\|_F^2 \tag{32}$$

where $\check{\mathbf{B}}_1^{t-1} := \rho/(2\rho+\eta)(\mathbf{B}_2^{t-1} + \mathbf{B}_3^{t-1} - \mathbf{M}_{12}^{t-1} - \mathbf{M}_{13}^{t-1}) + \eta/(2\rho+\eta)\mathbf{B}_1^{t-1} - \kappa_1/(2T(2\rho+\eta))\mathbf{P}$. Problem (32) could be reformulated and solved as a linearly-constrained quadratic program.

Interestingly enough though, the minimizer of (32) can be found in closed form for both choices of $f_{\boldsymbol{\pi}_t}(\mathbf{B}_1)$. Specifically, if $f_{\boldsymbol{\pi}_t}(\mathbf{B}_1) = \|\mathbf{B}_1\boldsymbol{\pi}_t\|_1$, the next claim is shown in the Appendix.

Proposition 2: Given $(\mathbf{Y}, \mathbf{z}) \in \mathbb{R}^{M\times N} \times \mathbb{R}^N$, the minimizer

$$\hat{\mathbf{X}} := \arg\min_{\mathbf{X}} \ \|\mathbf{X}\mathbf{z}\|_1 + \tfrac{1}{2}\|\mathbf{X} - \mathbf{Y}\|_F^2 \tag{33}$$

is given by $\hat{\mathbf{X}} = \mathbf{Y} - \mathcal{S}_z(\mathbf{Y}\mathbf{z})\mathbf{z}'$, where $z := \|\mathbf{z}\|_2^2$, and the operator $\mathcal{S}_z(\mathbf{x}) : \mathbb{R}^N \to \mathbb{R}^N$ is defined as $\mathcal{S}_z(x) := \operatorname{sign}(x) \cdot \min\{(|x|/z), 1\}$ applied entry-wise.

By Proposition 2, $\mathbf{B}_1^t$ can be found as a rank-one update of $\check{\mathbf{B}}_1^{t-1}$

$$\mathbf{B}_1^t = \check{\mathbf{B}}_1^{t-1} - \mathcal{S}_{\|\check{\boldsymbol{\pi}}_t\|_2^2}\Big(\check{\mathbf{B}}_1^{t-1}\check{\boldsymbol{\pi}}_t\Big)\check{\boldsymbol{\pi}}_t' \tag{34}$$

where $\check{\boldsymbol{\pi}}_t := 1/(2\rho+\eta)\boldsymbol{\pi}_t$. The key observation here is that having a single ℓ_1-norm $\|\mathbf{B}_1\boldsymbol{\pi}_t\|_1$ rather than $\|\mathbf{B}_1\mathbf{\Pi}\|_1 = \sum_{t=1}^{T}\|\mathbf{B}_1\boldsymbol{\pi}_t\|_1$ [see (11)] led to the simple update of (34).

For the Huber cost, the next claim is proved in the Appendix.

Algorithm 3 Topology Recovery Tracking Scheme

Require: $(\kappa_1, \kappa_2, \kappa_3, \rho, \eta)$.
1: Initialize $\mathbf{B}_1^0 = \mathbf{B}_2^0 = \mathbf{B}_3^0 = \mathbf{I}_N$.
2: Initialize $\mathbf{M}_{12}^0 = \mathbf{M}_{13}^0 = \mathbf{0}_N$.
3: **for** $t = 1, 2, \ldots, T$ **do**
4: Acquire price vector $\boldsymbol{\pi}_t$.
5: Update $\mathbf{B}_1^{t+1}$ using (34) or (37).
6: Update $(\mathbf{B}_2^{t+1}, \mathbf{B}_3^{t+1})$ from (38) and (39), respectively.
7: Update $(\mathbf{M}_{12}^{t+1}, \mathbf{M}_{13}^{t+1})$ from (40) and (41), respectively.
8: **end for**

Proposition 3: Given $(\mathbf{Y}, \mathbf{z}) \in \mathbb{R}^{M\times N} \times \mathbb{R}^N$, the minimizer

$$\hat{\mathbf{X}} := \arg\min_{\mathbf{X}} \ \alpha \tilde{h}_\kappa(\mathbf{X}\mathbf{z}) + \tfrac{1}{2}\|\mathbf{X} - \mathbf{Y}\|_F^2 \tag{35}$$

for $\alpha > 0$ is given by $\hat{\mathbf{X}} = \mathbf{Y} - \mathcal{H}_{z,\alpha,\kappa}(\mathbf{Y}\mathbf{z})\mathbf{z}'$, where $z := \|\mathbf{z}\|_2^2$, and $\mathcal{H}_{z,\alpha,\kappa}(\mathbf{x}) : \mathbb{R}^N \to \mathbb{R}^N$ is defined as

$$\mathcal{H}_{z,\alpha,\kappa}(x) := \begin{cases} \frac{x}{\alpha^{-1}+z}, & |x| \le \kappa(1+\alpha z) \\ \alpha\kappa, & x > \kappa(1+\alpha z) \\ -\alpha\kappa, & x < -\kappa(1+\alpha z) \end{cases} \tag{36}$$

applied entry-wise.

Based on Proposition 3, when $f_{\boldsymbol{\pi}_t}(\mathbf{B}_1) = \tilde{h}_{\kappa_3}(\mathbf{B}_1\boldsymbol{\pi}_t)$, the minimizer of (32) is

$$\mathbf{B}_1^t = \check{\mathbf{B}}_1^{t-1} - \mathcal{H}_{\|\boldsymbol{\pi}_t\|_2^2,(2\rho+\eta)^{-1},\kappa_3}\left(\check{\mathbf{B}}_1^{t-1}\boldsymbol{\pi}_t\right)\boldsymbol{\pi}_t'. \tag{37}$$

During the second step of iteration t, matrices $(\mathbf{B}_2, \mathbf{B}_3)$ are updated similarly to (20)–(23) as

$$\mathbf{B}_2^{t+1} = \min\left\{\mathbf{B}_1^{t+1} + \mathbf{M}_{12}^t, \mathbf{I}\right\} \tag{38}$$

$$\mathbf{B}_3^{t+1} = \mathcal{P}_{\kappa_2/(T\rho)}\left(\mathbf{B}_1^{t+1} + \mathbf{M}_{13}^t\right). \tag{39}$$

At the third step, the Lagrange multipliers are updated as

$$\mathbf{M}_{12}^{t+1} = \mathbf{M}_{12}^t + \mathbf{B}_1^{t+1} - \mathbf{B}_2^{t+1} \tag{40}$$

$$\mathbf{M}_{13}^{t+1} = \mathbf{M}_{13}^t + \mathbf{B}_1^{t+1} - \mathbf{B}_3^{t+1}. \tag{41}$$

To summarize, grid topology recovery using inexact streaming pricing data can be performed using Algorithm 3. Its complexity per time instance is dominated by the $\mathcal{O}(N^3)$ computations needed for the eigenvalue decomposition in (39).

VI. Experimental Validation

The computational complexity of the novel scheme was first verified using synthetic market data. Instances of the price matrix $\boldsymbol{\Pi}$ of dimension $N \times 5N$ were generated by independently drawing its entries from the standardized Gaussian distribution. All experiments were run on a 2.4 GHz Intel Core i7 (4GB RAM) laptop computer using MATLAB. The number of ADMM iterations and execution times for Algorithm 1 with $\kappa_1 = \kappa_2 = 0.5N$ are shown in Fig. 1. For N larger than 1000, execution times exceeded the level of hours, and thus, they were not tested here. In such cases, parallel computing, variable initialization, and custom-made routines for ADMM updates could be pursued. Alternatively, the online solvers tested later in this section become practically relevant.

To evaluate the recovery capabilities of the novel approaches, we then experimented with real-load data on the

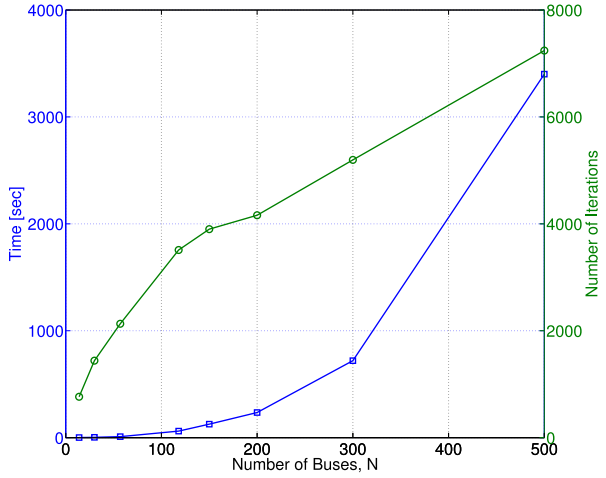

Fig. 1. Computational complexity of Algorithm 1 for price matrices $\boldsymbol{\Pi} \in \mathbb{R}^{N\times 5N}$.

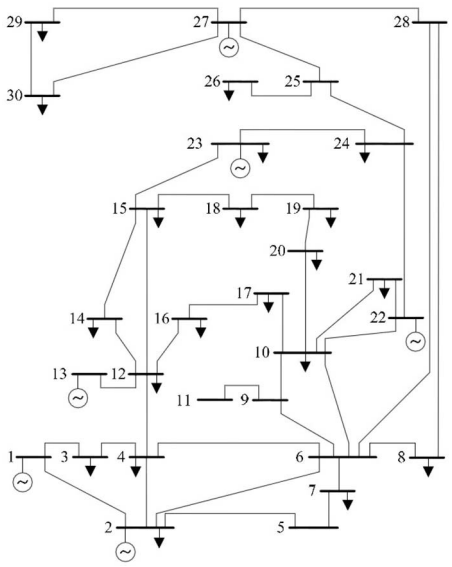

Fig. 2. Topology of the IEEE 30-bus system [30].

TABLE I
Generation Offers

Generator	Block offers [MWh,$/MWh]				
1	(30,26)	(20,36)	(20,44)	(10,50)	
2	(20,21)	(20,28)	(20,35)	(20,43)	
13	(15,38)	(15,42)	(10,47)		
22	(10,16)	(10,27)	(10,41)	(10,54)	(10,66)
23	(15,34)	(15,40)			
27	(30,35)	(15,39)			

IEEE 30-bus benchmark; see Fig. 2 [30]. The latter comprises 18 load buses, 6 generators, and 6 zero-injection buses. The transmission network consists of 41 lines with ratings ranging from 16 to 130 MVA as listed in [30].

Regarding offers, the benchmark provides generation capacities and quadratic generation costs [30]. To comply with market practices, the generation costs were first approximated by convex piece-wise linear functions yielding the block offers of Table I. The original costs were scaled up by ten to reflect current wholesale electricity cost levels. To model small

TABLE II
AVERAGE BUS DEGREE ATTAINED FOR (κ_1, κ_2)

$\kappa_1 \setminus \kappa_2$	0.001	0.01	0.1	1	10
0.001	0.9	1.2	1.5	2.5	5.9
0.01	0.9	1.0	1.4	2.5	5.9
0.1	0.8	1.0	1.6	2.4	5.9
1	0.4	2.8	1.9	2.7	5.9
10	6.5	5.6	5.9	6.0	6.0

fluctuations in costs, the nominal offers of Table I were shifted by a deviation uniformly distributed in $[-2.5, 2.5]$ \$/MWh.

For consumption, apart from the 18 load buses, generator buses 2 and 23 have load demands too, resulting in a total of 20 loads. The IEEE 30-bus benchmark provides a single realization of load demands. To simulate multiple realistic demands, we used the actual load data publicized for the global energy forecasting (GEF) competition 2012 [31]. These data are the hourly energy consumptions over 20 sites. To match the load levels of the IEEE 30-bus grid, all loads were scaled down by a factor of 7. The 20 demand sequences from December 23, 2007, were assigned to buses so that the average consumption per bus matched the demand specified by the benchmark. Hourly loads were perturbed by a zero-mean Gaussian variation having standard deviation ten times smaller than the nominal value to account for 5-min load fluctuations.

Real-time prices were generated by solving (2) for one day, i.e., 288 5-min intervals, and MCCs were announced separately. Lacking any day-ahead market information, the system was assumed to be dispatched entirely through the real-time market. Out of the 288 dispatches, three were infeasible and 45 experienced no congestion (occurred primarily over nighttime). Our experimental validation utilized the remaining $T = 240$ MCC price vectors. It is worth stressing that only lines (1,2), (15,23), and (6,28) became congested.

Upon collecting prices over an entire day, parameters (κ_1, κ_2) were selected. Although such parameters are typically tuned using cross-validation, this methodology becomes cumbersome for our problem. Assuming the average node degree for the grid of interest to be known, (κ_1, κ_2) were tuned so that the estimate $\hat{\mathbf{B}}$ had the same average degree. Given the scale ambiguity, the algorithm outcome $\hat{\mathbf{B}}$ was normalized by its maximum diagonal entry, and entries with absolute value smaller than 0.01 were set to zero.

Algorithm 1 was run for (κ_1, κ_2) taking the values $\{10^{-3}, 10^{-2}, 10^{-1}, 1, 10\}$. Regarding ρ, the convergence rate for the objective (constraint violation) is proportional (inversely proportional) to ρ [28]. For the problem at hand, setting $\rho = 10^4$ was empirically observed to provide a good trade-off. Since the average degree of the IEEE 30-bus grid is 2.68, the estimated node degrees obtained in Table II hint that (κ_1, κ_2) could be both set to 1. The actual and the recovered Laplacian matrix for the IEEE 30-bus benchmark are shown in Fig. 3.

We next simulated a longer observation interval: consumption data were generated by scaling GEF competition loads over January 2008 so that the maximum daily per-site value was 1.6 times the benchmark demands [31]. The precision-recall tradeoff of the topology recovery scheme was evaluated in terms of the true positive and false positive rates. The true

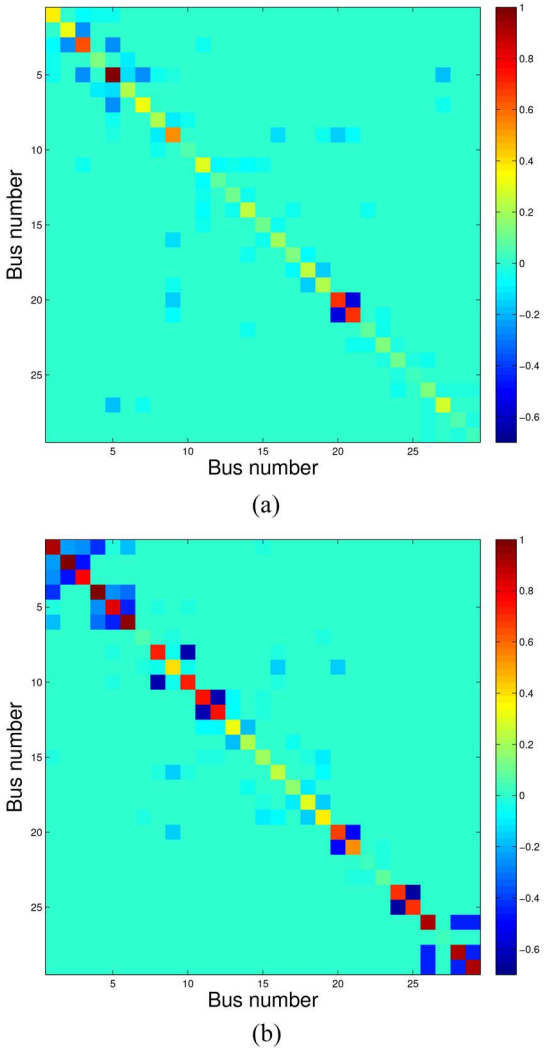

Fig. 3. Laplacian matrix for the IEEE 30-bus system. (a) Actual Laplacian matrix. (b) Laplacian matrix found by Algorithm 1 for $(\kappa_1, \kappa_2) = (1, 1)$.

positive rate was empirically calculated as the number of lines correctly identified over the total number of lines (39 for the reduced 30-bus grid). The false positive rate was measured as the ratio of transmission lines incorrectly identified as active over the number of bus pairs without a direct line connection (367 here). After ignoring market periods without congestion, Algorithm 1 was tested using successive observation intervals of varying lengths for different values of $\kappa_1 = \kappa_2$. Fig. 5 demonstrates that the precision-recall behavior of the scheme improves with increasing observation intervals as expected.

The method's tracking ability was tested by simulating a grid reconfiguration on January 15: lines (2,6) and (23,24) were exchanged for lines (2,7) and (23,26), respectively. Among the 8928 intervals, infeasible dispatches and dispatches without congestion were ignored yielding 7220 effective clearings. Algorithm 3 with the update of (37) was initialized to the batch solution obtained from Algorithm 1. Parameters ρ and η were set to $\sqrt{T}$ yielding sublinear regret [28], while κ_3 was set to 1. Fig. 4 depicts the tracking behavior of Algorithm 3. The estimated normalized reactance for line (10,17), i.e., entry $\hat{\mathbf{B}}_{9,16}$, remained relatively invariant. Line (2,7) was initially erroneously detected as active, yet it was adjusted

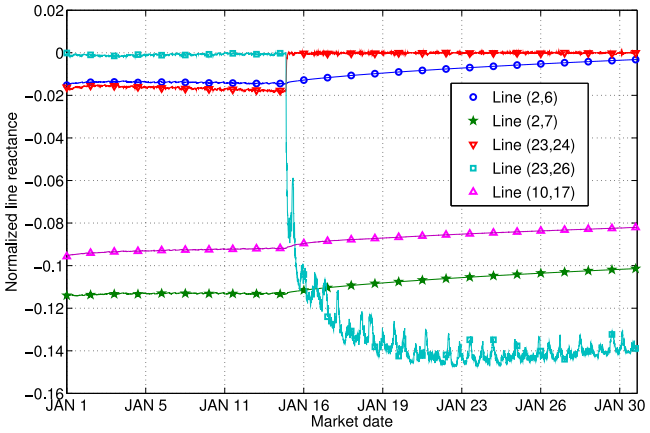

Fig. 4. Tracking lines using streaming pricing data for January 2008.

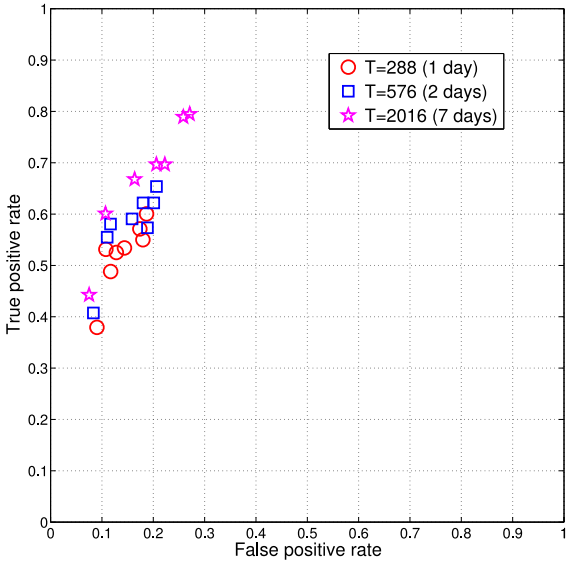

Fig. 5. Probability of correct detection over false alarms in identifying transmission lines for different observation intervals.

after January 15, while reactance (2,6) approached zero. Interestingly, the replacement of line (23,24) by (23,26) was promptly detected. Regarding computational complexity, one update needed 0.5 ms for $N = 30$ buses and scaled up to 5 s for $N = 2000$, deeming the online ADMM scheme pertinent for real-world energy markets.

VII. Conclusion

Grid topology recovery using publicly available energy prices was the subject of this paper. Upon exploiting the way real-time LMPs are obtained, recovery approaches with complementary strengths were developed. Advances in compressive sampling and online convex optimization proved to be useful for grid topology tracking. Experimental validation using real-consumption data on a benchmark grid corroborated the risk of unveiling the power network structure. Numerical tests using a month-long price dataset showed the possibility of tracking grid reconfigurations too. The recovery performance could be enhanced further in envisioned smart grids. In competitive markets, rapidly changing offers and bids could probe the dispatch LP in a richer way, while market data announced at higher rates could provide even more information. Characterizing identifiability is a challenging yet interesting research direction. Moreover, leveraging heterogeneous market data, such as prior knowledge on transmission lines, local congestion information and schedules, and historical global congestion and schedule information are expected to improve the recovery results.

Appendix

Proof of Proposition 2: Strict convexity of $(1/2)\|\mathbf{X}-\mathbf{Y}\|_F^2$ implies that (33) admits a unique minimizer. First-order optimality conditions assert that $\mathbf{0}$ belongs to the subdifferential of $\|\mathbf{Xz}\|_1 + (1/2)\|\mathbf{X}-\mathbf{Y}\|_F^2$ evaluated at $\hat{\mathbf{X}}$. By definition, the subdifferential of $\|\mathbf{Xz}\|_1$ at $\hat{\mathbf{X}}$ is $\hat{\mathbf{g}}\mathbf{z}'$, where

$$\hat{g}_n' := \begin{cases} \text{sign}(\hat{\mathbf{x}}_n'\mathbf{z}), & \hat{\mathbf{x}}_n'\mathbf{z} \neq 0 \\ s_n : |s_n| \leq 1, & \text{otherwise} \end{cases} \tag{42}$$

is the n-th entry of $\hat{\mathbf{g}}$, and $\hat{\mathbf{x}}_n'$ denotes the n-th row of $\hat{\mathbf{X}}$. Hence, the first-order optimality condition implies that

$$\hat{\mathbf{X}} = \mathbf{Y} - \hat{\mathbf{g}}\mathbf{z}'. \tag{43}$$

Unless $\mathbf{z} = \mathbf{0}$ and trivially $\hat{\mathbf{X}} = \mathbf{Y}$, the minimizer $\hat{\mathbf{X}}$ is a rank-one update of $\mathbf{Y}$ granted $\hat{\mathbf{g}}$ is known. To find $\hat{\mathbf{g}}$, post-multiply (43) by $\mathbf{z}$ to get $\hat{\mathbf{X}}\mathbf{z} = \mathbf{Y}\mathbf{z} - \hat{\mathbf{g}}z$ whose n-th entry reads

$$\hat{\mathbf{x}}_n'\mathbf{z} = \mathbf{y}_n'\mathbf{z} - \hat{g}_n z \tag{44}$$

with $z := \|\mathbf{z}\|_2^2$ and $\mathbf{y}_n$ being the n-th row of $\mathbf{Y}$. Given (42) and depending on $\mathbf{y}_n'\mathbf{z}$, three cases can be identified for (44): **(c1)** If $\mathbf{y}_n'\mathbf{z} > z$, then $\hat{g}_n = +1$ and $\hat{\mathbf{x}}_n'\mathbf{z} > 0$; **(c2)** if $\mathbf{y}_n'\mathbf{z} < -z$, then $\hat{g}_n = -1$ and $\hat{\mathbf{x}}_n'\mathbf{z} < 0$; and **(c3)** if $|\mathbf{y}_n'\mathbf{z}| \leq z$, then $\hat{g}_n = \mathbf{y}_n'\mathbf{z}/z$ and $\hat{\mathbf{x}}_n'\mathbf{z} = 0$; thus proving the claim. ■

Proof of Proposition 3: Similar to Proposition 2, first-order optimality conditions imply that

$$\hat{\mathbf{X}} = \mathbf{Y} - \alpha\hat{\mathbf{g}}\mathbf{z}' \tag{45}$$

where the n-th entry of $\hat{\mathbf{g}}$ is defined as

$$\hat{g}_n' := \begin{cases} \hat{\mathbf{x}}_n'\mathbf{z}, & |\hat{\mathbf{x}}_n'\mathbf{z}| \leq \kappa \\ \kappa\,\text{sign}(\hat{\mathbf{x}}_n'\mathbf{z}), & \text{otherwise} \end{cases} \tag{46}$$

and $\hat{\mathbf{x}}_n'$ is the n-th row of $\hat{\mathbf{X}}$. To find $\hat{\mathbf{g}}$, post-multiply (45) by $\mathbf{z}$ to obtain $\hat{\mathbf{X}}\mathbf{z} = \mathbf{Y}\mathbf{z} - \alpha\hat{\mathbf{g}}z$, whose n-th entry reads

$$\hat{\mathbf{x}}_n'\mathbf{z} = \mathbf{y}_n'\mathbf{z} - \alpha\hat{g}_n z \tag{47}$$

with $z := \|\mathbf{z}\|_2^2$ and $\mathbf{y}_n$ being the n-th row of $\mathbf{Y}$. Based on (46), three cases can be distinguished for (47): **(c1)** If $|\mathbf{y}_n'\mathbf{z}| \leq \kappa(1+\alpha z)$, then $\hat{g}_n = \hat{\mathbf{x}}_n'\mathbf{z}$; **(c2)** if $\mathbf{y}_n'\mathbf{z} > \kappa(1+\alpha z)$, then $\hat{g}_n = \kappa$; and **(c3)** if $\mathbf{y}_n'\mathbf{z} < -\kappa(1+\alpha z)$, then $\hat{g}_n = -\kappa$. Note that for (c1), $\hat{g}_n$ depends on the unknown $\hat{\mathbf{x}}_n$. By substituting $\hat{g}_n$ back into (45) and focusing on the n-th row of $\hat{\mathbf{X}}$, we arrive at $(\mathbf{I}+\alpha\mathbf{z}\mathbf{z}')\hat{\mathbf{x}}_n = \mathbf{y}_n$. Invoking the matrix inversion lemma yields $\hat{\mathbf{x}}_n = \mathbf{y}_n - (\mathbf{y}_n'\mathbf{z}/(\alpha^{-1}+z))\mathbf{z}$. ■

References

[1] A. L. Ott, "Experience with PJM market operation, system design, and implementation," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 528–534, May 2003.

[2] *Smart Grid Research and Development, Multi-Year Program Plan (2010–2014)*, U.S. Dept. Energy, Washington, DC, USA, Sep. 2011.

[3] G. B. Giannakis *et al.*, "Monitoring and optimization for power grids: A signal processing perspective," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 107–128, Sep. 2013.
[4] A. Abur and A. Gómez-Expósito, *Power System State Estimation: Theory and Implementation*. New York, NY, USA: Marcel Dekker, 2004.
[5] R. Emami and A. Abur, "Tracking changes in the external network model," in *Proc. North Amer. Power Symp.*, Arlington, TX, USA, Sep. 2010, pp. 1–6.
[6] J. E. Tate and T. J. Overbye, "Double line outage detection using phasor angle measurements," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Calgary, AB, Canada, Jul. 2009, pp. 1–5.
[7] H. Zhu and G. B. Giannakis, "Sparse overcomplete representations for efficient identification of power line outages," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2215–2224, Nov. 2012.
[8] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Trans. Inf. Syst. Security*, vol. 14, no. 1, pp. 1–33, May 2011.
[9] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 645–658, Dec. 2011.
[10] L. Jia, J. Kim, R. J. Thomas, and L. Tong, "Impact of data quality on real-time locational marginal price," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 627–636, Mar. 2014.
[11] J. Kim, L. Tong, and R. J. Thomas, "Data framing attack on state estimation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 7, pp. 1460–1470, Jul. 2014.
[12] F. Pasqualetti, F. Dorfler, and F. Bullo, "Cyber-physical attacks in power networks: Models, fundamental limitations and monitor design," in *Proc. IEEE Conf. Decis. Control*, Orlando, FL, USA, Dec. 2011, pp. 2195–2201.
[13] K. C. Sou, H. Sandberg, and K. H. Johansson, "Data attack isolation in power networks using secure voltage magnitude measurements," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 14–28, Jan. 2014.
[14] L. Xie, Y. Mo, and B. Sinopoli, "Integrity data attacks in power market operations," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 659–666, Dec. 2011.
[15] W. Xu, M. Wang, J.-F. Cai, and A. Tang, "Sparse error correction from nonlinear measurements with applications in bad data detection for power networks," *IEEE Trans. Signal Process.*, vol. 61, no. 24, pp. 6175–6187, Dec. 2013.
[16] Y.-Y. Lee, J. Hur, R. Baldick, and S. Pineda, "New indices of market power in transmission-constrained electricity markets," *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 681–689, May 2011.
[17] V. Kekatos, Y. Zhang, and G. B. Giannakis, "Electricity market forecasting via low-rank multi-kernel learning," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 6, pp. 1182–1193, Dec. 2014.
[18] X. Li, H. V. Poor, and A. Scaglione, "Blind topology identification for power systems," in *Proc. IEEE Smart Grid Commun. Conf.*, Vancouver, BC, Canada, Oct. 2013, pp. 91–96.
[19] T. Erseghe, S. Tomasin, and A. Vigato, "Topology estimation for smart micro grids via powerline communications," *IEEE Trans. Signal Process.*, vol. 61, no. 13, pp. 3368–3377, Jul. 2013.
[20] M. He and J. Zhang, "A dependency graph approach for fault detection and localization towards secure smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 2, pp. 342–351, Jun. 2011.
[21] S. Bolognani, N. Bof, D. Michelotti, R. Muraro, and L. Schenato, "Identification of power distribution network topology via voltage correlation analysis," in *Proc. IEEE Conf. Decis. Control*, Florence, Italy, Dec. 2013, pp. 1659–1664.
[22] V. Kekatos, G. B. Giannakis, and R. Baldick, "Grid topology identification using electricity prices," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, National Harbor, MD, USA, Jul. 2014, pp. 1–5.
[23] A. Gómez-Expósito, A. J. Conejo, and C. Canizares, Eds., *Electric Energy Systems, Analysis and Operation*. Boca Raton, FL, USA: CRC Press, 2009.
[24] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1991.
[25] J. E. Price and J. Goodin, "Reduced network modeling of WECC as a market design prototype," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, San Diego, CA, USA, Jul. 2011, pp. 1–6.
[26] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM J. Comput.*, vol. 24, no. 2, pp. 227–234, 1995.
[27] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
[28] H. Wang and A. Banerjee, "Online alternating direction method," in *Proc. Int. Conf. Mach. Learn.*, Edinburgh, U.K., Jul. 2012, pp. 1–8.
[29] J. Duchi, S. Shalev-Shwartz, Y. Singer, and A. Tewari, "Composite objective mirror descent," in *Proc. Conf. Learn. Theory*, Haifa, Israel, Jun. 2010, pp. 1–13.
[30] *Power Systems Test Case Archive. University of Washington*. [Online]. Available: http://www.ee.washington.edu/research/pstca/, accessed Aug. 22, 2015.
[31] *Load Forecasting. Global Energy Forecasting Competition 2012*. [Online]. Available: https://www.kaggle.com/c/global-energy-forecasting-competition-2012-load-forecasting, accessed Aug. 22, 2015.

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

He was a Research Associate with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN, USA. In 2012, he was with Windlogics Inc., Saint Paul, MN. In 2014, he was a Visiting Researcher with the University of Texas at Austin, Austin, TX, USA, and Ohio State University, Columbus, OH, USA. In 2015, he joined the Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA, USA, as an Assistant Professor. His current research interests include optimization, learning, and management of future energy systems.

Dr. Kekatos was a recipient of the Marie Curie Fellowship and the Postdoctoral Career Development Award (honorable mention) from the University of Minnesota, in 2014.

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

Since 1999, he has been a Professor with the University of Minnesota, Minneapolis, MN, USA, where he is currently the ADC Chair of Wireless Telecommunications, Department of Electrical Communication Engineering, and serves as the Director of the Digital Technology Center. His current research interests include communications, networking, and statistical signal processing subjects; sparsity and big data analytics; wireless cognitive radios; mobile ad hoc networks; renewable energy; power grids; gene-regulatory; and social networks. He has published over 380 journal papers, 650 conference papers, 20 book chapters, two edited books, and two research monographs with an h-index of 113. He is a (co-)inventor of 22 patents issued.

Prof. Giannakis was a (co-)recipient of eight best paper awards from the IEEE Signal Processing (SP) and Communications Societies, including the G. Marconi Prize Paper Award in Wireless Communications, the Technical Achievement Awards from the SP Society in 2000 and European Association for Signal Processing (EURASIP) in 2005, the Young Faculty Teaching Award, the G. W. Taylor Award for Distinguished Research from the University of Minnesota, and the IEEE Fourier Technical Field Award in 2014. He has served the IEEE in a number of posts, including a Distinguished Lecturer for the IEEE SP Society. He is a Fellow of EURASIP.

Ross Baldick (F'07) received the B.Sc. degree in mathematics and physics, and the B.E. degree in electrical engineering from the University of Sydney, Sydney, NSW, Australia, and the M.S. and Ph.D. degrees in electrical engineering and computer sciences from the University of California–Berkeley, Berkeley, CA, USA, in 1988 and 1990, respectively.

From 1991 to 1992, he was a Postdoctoral Fellow with Lawrence Berkeley Laboratory, Berkeley. In 1992 and 1993, he was an Assistant Professor with Worcester Polytechnic Institute, Worcester, MA, USA. He is currently a Professor with the Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX, USA.