Digital Object Identifier: 10.1109/TPWRS.2013.2293634

Sensitivity Analysis of Real-Time Locational Marginal Price to SCADA Sensor Data Corruption

Dae-Hyun Choi, *Student Member, IEEE*, and Le Xie, *Member, IEEE*

***Abstract*—This paper examines the impact of supervisory control and data acquisition (SCADA) data corruption on real-time locational marginal price (LMP) in electricity markets. We present an analytical framework to quantify LMP sensitivity with respect to changes in sensor data. This framework consists of a unified LMP sensitivity matrix subject to sensor data corruption. This sensitivity matrix reflects a coupling among the sensor data, an estimation of the power system states, and the real-time LMP. The proposed framework offers system operators an online tool to: 1) quantify the impact of corrupted data at any sensor on LMP variations at any bus; 2) identify buses with LMPs highly sensitive to data corruption; and 3) find sensors that impact LMP changes significantly and influentially. It also allows system operators to evaluate the impact of SCADA data accuracy on real-time LMP. The results of the proposed sensitivity based analysis are illustrated and verified with IEEE 14-bus and 118-bus systems with both Ex-ante and Ex-post real-time pricing models.**

***Index Terms*—Economic dispatch, locational marginal price (LMP), power market, supervisory control and data acquisition (SCADA), sensitivity analysis, state estimation.**

NOMENCLATURE

The main notations used throughout this paper are summarized here. Bold symbols represent vectors or matrices. Hat symbols represent estimates of true parameter value.

a_i	Linear cost coefficient for generator i.
b_i	Quadratic cost coefficient for generator i.
$C_i(\cdot)$	Energy cost for generator i.
P_{g_i}	Scheduled generator power output for generator i.
L_{d_i}	Fixed demand at bus i.
$P_{g_i}^{\min}, P_{g_i}^{\max}$	Min/max generation limits for generator i at Ex-ante dispatch.
$F_l^{\min}, F_l^{\max}$	Min/max flow limits for transmission line l at Ex-ante dispatch.
S_{li}	Generation shift factor of transmission line l to bus i.
$\Delta P_{g_i}^{\max}, \Delta P_{g_i}^{\min}$	Min/max incremental generation limits for generator i at Ex-post dispatch.
R_i	Ramp rate of generator i.
ΔT	Dispatch interval.
π_i	Locational marginal price at bus i.
λ	Shadow price of the system energy balance equation.
τ_i	Shadow price of the capacity constraint for generator i.
μ_l	Shadow price of the transmission line constraint for transmission line l.
N_b	Total number of buses.
N_m	Total number of sensor measurements.
N_l	Total number of transmission lines.
$\mathcal{CL}_+, \mathcal{CL}_-$	Sets of positively and negatively congested lines at Ex-ante dispatch.
$\mathcal{S}_v$	Set of voltage magnitude measurements.
$\mathcal{S}_{ri}$	Set of real power injection measurements.
$\mathcal{S}_{ai}$	Set of reactive power injection measurements.
$\mathcal{S}_{rf}$	Set of real power flow measurements.
$\mathcal{S}_{af}$	Set of reactive power flow measurements.
$\mathbf{I}_k$	$k \times k$ identity matrix.
$\mathbf{1}_k, \mathbf{0}_k$	$k \times 1$ column vectors with all ones and all zeros, respectively.

I. INTRODUCTION

STATE estimation is one of the key applications for power-system energy management systems (EMSs). The impact of bad data on power systems has been intensively investigated in recent decades in power-system state estimation literature [1]. Measurement noise and/or manipulated sensor errors in a supervisory control and data acquisition (SCADA) system may mislead system operators about real-time conditions in a power system, which in turn may impact the price signals in real-time power markets. This paper attempts to provide a novel analytical framework with which to investigate the impact of

Manuscript received February 15, 2013; revised June 22, 2013 and September 04, 2013; accepted November 13, 2013. This work was supported in part by the Power Systems Engineering Research Center, the National Science Foundation under DUE Grant 1303378, and the National Science Foundation under ECCS Grant 1150944. Paper no. TPWRS-00192-2013.

The authors are with the Department of Electrical and Computer Engineering, Texas A&M University, College Station TX, USA (e-mail: cdh8954@neo.tamu.edu; Lxie@ece.tamu.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.2013.2293634

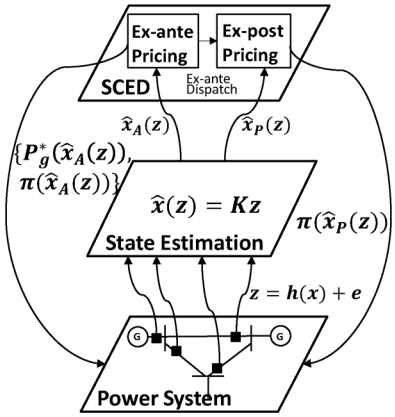

Fig. 1. Three-layered framework illustrating the coupling of the physical power system, state estimation, and SCED.

bad sensor data on electric *power market* operations. In future power-system operations, which will probably involve many more sensors, the impact of sensor data quality on grid operations will become increasingly important.

Locational marginal price (LMP) is the core variable in market operations [2]. In real-time power markets, LMP is obtained as the by-product of security-constrained economic dispatch (SCED) in either of the two main pricing models: Ex-ante (e.g., in ERCOT, NY ISO) and Ex-post (e.g., in ISO New England, PJM, and Midwest ISO) [3]. Both pricing models are built on the power flow and network topology results given by the state estimator, which uses two types of sensor data: 1) analog (e.g., the power injection/flow and voltage magnitude) and 2) digital (e.g., the on/off status of a circuit breaker). In this paper, we focus on a sensitivity analysis of real-time LMP subject to corrupted *analog* data fed into the state estimator. Fig. 1 illustrates that, via state estimation, SCADA measurement $\mathbf{z}$ may impact the results of a pair of Ex-ante nodal price and optimal generation dispatch $\{\boldsymbol{\pi}(\hat{\mathbf{x}}_A(\mathbf{z})),\ P_g^*(\hat{\mathbf{x}}_A(\mathbf{z}))\}$ and the Ex-post price $\boldsymbol{\pi}(\hat{\mathbf{x}}_P(\mathbf{z}))$.

Real-time market LMPs are primarily affected by a system's physical conditions, which are the results of state estimation routine. A study of LMP sensitivity with respect to system physical conditions was first conducted by Conejo *et al.* [4]. In this work, the LMP sensitivity problem was formulated in nonlinear programming based on the ac optimal power flow (ACOPF) model. It provided a generalized platform for calculating the sensitivity of LMP with respect to changes in various parameters such as load, generator cost, voltage limit, generation power limit, and network topology. Sensitivity studies have also been conducted with linear programming based on the dc optimal power flow (DCOPF) model with a DCOPF-based algorithm [5], the probabilistic model [6], and the continuous locational marginal pricing approach [7]. All previous work has focused mainly on the impact of physical load variations on LMP sensitivity. More recently, some work has studied the economic impact of cyber data attacks on real-time power market operations. This recent work has demonstrated that data corruption from an adversary can bypass the Chi-squares bad data detection [8], [9], consequently leading to LMP distortion due to state estimation error [10], [11]. Our paper is motivated by a desire to study the effect of data corruption on LMP via *state estimation*. In particular, this paper provides an analytical framework for answering the following questions.

1) How much does LMP change at every bus given a set of SCADA measurements with corrupted data?
2) What is the impact of data accuracy on LMP sensitivity at each bus?

Here, data corruption refers to both natural noise and man-made attacks.

The novel aspect of this paper is that it provides system operators with an analytical tool for assessing the financial risks of bad/malicious data in light of secure market operations. To this end, a unified LMP sensitivity matrix subject to data corruption is developed, describing the coupling between LMP, the estimation of power system states, and the sensor data. This matrix offers system operators an *online* tool to: 1) quantify the impact of corrupted data at any sensor on LMP variation at any bus; 2) identify buses with LMP highly sensitive to data corruption; 3) find significant and influential sensors with regards to LMP change; and 4) study the effect of data accuracy on LMP sensitivity.

The remainder of this paper is organized as follows. We briefly review state estimation and two representative real-time pricing models in Section II. We formulate the problem in Section III and derive the quantifying sensitivity of LMP subject to corrupted data in Section IV. Section V presents numerical examples that illustrate the impact of different SCADA sensors on LMP in IEEE 14-bus and 118-bus systems with both the Ex-ante and Ex-post pricing models. We make concluding remarks and suggest future work in Section VII.

II. Preliminaries

A. State Estimation Model

The measurement model for state estimation is formulated as

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \mathbf{e} \tag{1}$$

where $\mathbf{z} = [\mathbf{z}_r^T \mathbf{z}_a^T \mathbf{z}_v^T]^T$ is the $N_m \times 1$ measurement vector that consists of real power injection and the flow vector $\mathbf{z}_r = [\mathbf{z}_{ri}^T \mathbf{z}_{rf}^T]^T$, the reactive power injection and flow vector $\mathbf{z}_a = [\mathbf{z}_{ai}^T \mathbf{z}_{af}^T]^T$, and the bus voltage magnitude vector $\mathbf{z}_v$. $\mathbf{x} = [\boldsymbol{\theta}^T \mathbf{V}^T]^T$ is the state vector that consists of the $(N_b - 1) \times 1$ bus voltage phase angle vector $\boldsymbol{\theta}$ excluding a slack bus and the $N_b \times 1$ voltage magnitude vector $\mathbf{V}$. $\mathbf{h}(\mathbf{x})$ is the $N_m \times 1$ nonlinear vector valued measurement function relating measurements to states, and $\mathbf{e}$ is the $N_m \times 1$ independent identically distributed (i.i.d.) Gaussian measurement error vector with zero mean and diagonal covariance matrix $\mathbf{R}$. The state estimator computes the optimal estimate of $\mathbf{x}$ by minimizing the weighted least squares of measurement error as follows:

$$\text{minimize } J(\mathbf{x}) = \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r} \tag{2}$$

$$\text{s.t. } \mathbf{r} = \mathbf{z} - \mathbf{h}(\mathbf{x}). \tag{3}$$

Using the Gauss–Newton method, the weighted least-squares estimate vector $\hat{\mathbf{x}}$ is computed by the following iterative procedure [12]:

$$\Delta\hat{\mathbf{x}}^{k+1} = \left[\mathbf{G}(\hat{\mathbf{x}}^k)\right]^{-1}\mathbf{H}^T(\hat{\mathbf{x}}^k)\mathbf{R}^{-1}\Delta\mathbf{z}^k \tag{4}$$

where $\mathbf{H}(\hat{\mathbf{x}}^k) = [\partial\mathbf{h}(\hat{\mathbf{x}}^k)/\partial\hat{\mathbf{x}}^k]$ is the $N_m \times (2N_b - 1)$ Jacobian matrix at kth iteration, and

$$\Delta\hat{\mathbf{x}}^{k+1} = \hat{\mathbf{x}}^{k+1} - \hat{\mathbf{x}}^k \tag{5}$$
$$\Delta\mathbf{z}^k = \mathbf{z} - \mathbf{h}(\hat{\mathbf{x}}^k) \tag{6}$$
$$\mathbf{G}(\hat{\mathbf{x}}^k) = \mathbf{H}^T(\hat{\mathbf{x}}^k)\mathbf{R}^{-1}\mathbf{H}(\hat{\mathbf{x}}^k). \tag{7}$$

The iteration process in (4) continues until the maximum of $|\Delta\hat{\mathbf{x}}^k|$ is less than a predetermined threshold, otherwise it stops and yields the ultimate estimates.

B. Real-Time Electricity Pricing Model

State estimation results impact real-time market operations. Here, we present two real-time pricing models based on the DCOPF model. For simplicity, we assume that each bus has one generator and one load.

1) Ex-Ante Model: In ex-ante real-time market models, LMPs are computed before the actual deployment of dispatch orders. For the system operator, the Ex-ante dispatch is formulated as follows [5]:

$$\min_{P_{g_i}} \sum_{i=1}^{N_b} C_i\left(P_{g_i}\right) \tag{8}$$

s.t.

$$\lambda : \sum_{i=1}^{N_b} P_{g_i} = \sum_{i=1}^{N_b} L_{d_i} \tag{9}$$
$$\boldsymbol{\tau} : \hat{P}_{g_i}^{\min} \le P_{g_i} \le \hat{P}_{g_i}^{\max} \quad \forall i = 1, \ldots, N_b \tag{10}$$
$$\boldsymbol{\mu} : F_l^{\min} \le \sum_{i=1}^{N_b} S_{li}\left(P_{g_i} - L_{d_i}\right) \le F_l^{\max} \,\forall l = 1, \ldots, N_l \tag{11}$$

where

$$\hat{P}_{g_i}^{\max} = \min\left\{P_{g_i}^{\max}, \hat{P}_{g_i}(\mathbf{z}) + R_i\Delta T\right\}$$
$$\hat{P}_{g_i}^{\min} = \max\left\{P_{g_i}^{\min}, \hat{P}_{g_i}(\mathbf{z}) - R_i\Delta T\right\}.$$

In this formulation, the objective function is to minimize the total generation costs in (8). Equation (9) is the system-wide energy balance equation, (10) is the physical capacity constraints of each generator embedded with its ramp constraints, and (11) is the transmission line constraints.

2) Ex-Post Model: In ex-post real-time market models, LMPs are computed after the fact using real-time estimates for settlement purposes. Assuming no demand elasticity, the Ex-post dispatch is written as [13]

$$\min_{P_{g_i}} \sum_{i=1}^{N_b} C_i\left(P_{g_i}\right) \tag{12}$$

s.t.

$$\lambda : \sum_{i=1}^{N_b} P_{g_i} = \sum_{i=1}^{N_b} \hat{P}_{g_i}(\mathbf{z}) \tag{13}$$
$$\boldsymbol{\tau} : \hat{P}_{g_i}^{\min} \le P_{g_i} \le \hat{P}_{g_i}^{\max} \quad \forall i = 1, \ldots, N_b \tag{14}$$
$$\boldsymbol{\mu}_{\max} : \sum_{i=1}^{N_b} S_{li}\left(P_{g_i} - L_{d_i}\right) \le \hat{F}_l(\mathbf{z}) \,\forall l \in \mathcal{CL}_+ \tag{15}$$
$$\boldsymbol{\mu}_{\min} : \sum_{i=1}^{N_b} S_{li}\left(P_{g_i} - L_{d_i}\right) \ge \hat{F}_l(\mathbf{z}) \,\forall l \in \mathcal{CL}_- \tag{16}$$

where

$$\hat{P}_{g_i}^{\max} = \hat{P}_{g_i}(\mathbf{z}) + \Delta P_{g_i}^{\max}, \hat{P}_{g_i}^{\min} = \hat{P}_{g_i}(\mathbf{z}) + \Delta P_{g_i}^{\min}.$$

The above formulation is expressed with different notation than the Ex-post model formulated in [13] in order to emphasize that the state estimation solution has a direct impact on the Ex-post model.

III. Problem Formulation

Referring to Fig. 1, for all buses $(i = 1, \ldots, N_b)$ and measurements $(j = 1, \ldots, N_m)$, the $N_b \times 1$ vector of LMPs can be expressed in a composite function form

$$\mathbf{LMP} = \boldsymbol{\pi}\left(\hat{\mathbf{x}}(\mathbf{z})\right)$$

where

$$\boldsymbol{\pi} = \left[\pi_1, \pi_2, \ldots, \pi_{N_b}\right]^T \tag{17}$$
$$\pi_i = f_i\left(\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_{N_m}\right) \tag{18}$$
$$\hat{x}_j = g_j\left(z_1, z_2, \cdots, z_{N_m}\right). \tag{19}$$

π_i represents the LMP at bus i. z_j and $\hat{x}_j$ are the measurement and its corresponding estimate at sensor j, respectively. $f_i(\cdot)$ is the vector function that describes the relationship between any estimate and LMP at bus i, and $g_j(\cdot)$ is the vector function that describes the relationship between any measurement and estimate at sensor j.

The primary goal of this paper is to compute LMP sensitivity at any bus i subject to a measurement change at any sensor j throughout the entire transmission network:

$$\frac{\partial\pi_i}{\partial z_j} = \mathbf{\Lambda}_{(i,j)}. \tag{20}$$

By the chain rule, for all i and j, (20) is written as

$$\frac{\partial\pi_i}{\partial z_j} = \frac{\partial\pi_i}{\partial\hat{x}_1}\frac{\partial\hat{x}_1}{\partial z_j} + \frac{\partial\pi_i}{\partial\hat{x}_2}\frac{\partial\hat{x}_2}{\partial z_j} + \cdots + \frac{\partial\pi_i}{\partial\hat{x}_{N_m}}\frac{\partial\hat{x}_{N_m}}{\partial z_j}. \tag{21}$$

In (21), the estimate $\hat{x}_j$ is chosen as an intermediate variable for computing the partial derivative of π_i with respect to z_j. This variable is used to set the bounds for: 1) minimum and maximum generation capacity in (10) and (14); 2) the system balance equation in (13); and 3) the positive and negative transmission line capacity in (15) and (16). Equation (21) can be expressed in matrix form as shown as

$$\begin{aligned}\mathbf{\Lambda}_{(N_b\times N_m)} &= \frac{\partial \boldsymbol{\pi}}{\partial \mathbf{z}} = \frac{\partial \boldsymbol{\pi}}{\partial \hat{\mathbf{x}}}\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{z}} \\ &= \underbrace{\begin{bmatrix} \frac{\partial \pi_1}{\partial \hat{x}_1} & \frac{\partial \pi_1}{\partial \hat{x}_2} & \cdots & \frac{\partial \pi_1}{\partial \hat{x}_{N_m}} \\ \frac{\partial \pi_2}{\partial \hat{x}_1} & \frac{\partial \pi_2}{\partial \hat{x}_2} & \cdots & \frac{\partial \pi_2}{\partial \hat{x}_{N_m}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \pi_{N_b}}{\partial \hat{x}_1} & \frac{\partial \pi_{N_b}}{\partial \hat{x}_2} & \cdots & \frac{\partial \pi_{N_b}}{\partial \hat{x}_{N_m}} \end{bmatrix}}_{\mathbf{\Lambda}_A} \\ &\quad \times \underbrace{\begin{bmatrix} \frac{\partial \hat{x}_1}{\partial z_1} & \frac{\partial \hat{x}_1}{\partial z_2} & \cdots & \frac{\partial \hat{x}_1}{\partial z_{N_m}} \\ \frac{\partial \hat{x}_2}{\partial z_1} & \frac{\partial \hat{x}_2}{\partial z_2} & \cdots & \frac{\partial \hat{x}_2}{\partial z_{N_m}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \hat{x}_{N_m}}{\partial z_1} & \frac{\partial \hat{x}_{N_m}}{\partial z_2} & \cdots & \frac{\partial \hat{x}_{N_m}}{\partial z_{N_m}} \end{bmatrix}}_{\mathbf{\Lambda}_B}. \end{aligned} \quad (22)$$

The sensitivity $\mathbf{\Lambda}_{(i,j)}$ in (20) is the element at the ith row and jth column of the $N_b \times N_m$ sensitivity matrix $\mathbf{\Lambda}$. The matrix $\mathbf{\Lambda}$ is written as the multiplication form of two matrices with different types of sensitivities: the $N_b \times N_m$ matrix $\mathbf{\Lambda}_A = \partial\boldsymbol{\pi}/\partial\hat{\mathbf{x}}$ quantifies the *economic* impact of any estimate on any LMP, and the $N_m \times N_m$ matrix $\mathbf{\Lambda}_B = \partial\hat{\mathbf{x}}/\partial\mathbf{z}$ quantifies the *cyber* impact of any sensor measurement on any estimate. The derivations of $\mathbf{\Lambda}_A$ and $\mathbf{\Lambda}_B$ are described in more detail in the next section.

IV. LMP Sensitivity to Sensor Data Corruption

A. Sensitivity of LMPs to Estimated States

We first derive the sensitivity matrix $\mathbf{\Lambda}_A$ using the Ex-ante model. To this end, the perturbation approach developed in [4] is applied to the Ex-ante model in Section II-B. The Lagrangian function of the Ex-ante dispatch is written as

$$\begin{aligned}\mathcal{L} = &\sum_{i=1}^{N_b} C_i\left(P_{g_i}\right) - \lambda\left(\sum_{i=1}^{N_b}\left[P_{g_i} - L_{d_i}\right]\right) \\ &+ \sum_{j=1}^{2N_b}\tau_j\left(\sum_{i=1}^{N_b} A_{ji}P_{g_i} - \hat{C}_j\right) \\ &+ \sum_{l=1}^{2N_l}\mu_l\left(\sum_{i=1}^{N_b} S_{li}[P_{g_i} - L_{d_i}] - D_l\right)\end{aligned}$$

where A_{ji}, S_{li}, $\hat{C}_j$ and D_l are the elements of the following matrices:

$$\mathbf{A}_{(2N_b\times N_b)} = [A_{ji}] = \begin{bmatrix}\mathbf{I}_{N_b} \\ -\mathbf{I}_{N_b}\end{bmatrix} \quad (23)$$

$$\mathbf{B}_{(2N_l\times N_b)} = [S_{li}] = \begin{bmatrix}\mathbf{S} \\ -\mathbf{S}\end{bmatrix} \quad (24)$$

$$\hat{\mathbf{C}}_{(2N_b\times 1)} = [\hat{C}_j] = \begin{bmatrix}\hat{\mathbf{P}}_g^{\max} \\ -\hat{\mathbf{P}}_g^{\min}\end{bmatrix} \quad (25)$$

$$\mathbf{D}_{(2N_l\times 1)} = [D_l] = \begin{bmatrix}\mathbf{F}^{\max} \\ -\mathbf{F}^{\min}\end{bmatrix} \quad (26)$$

where $\mathbf{S}$ is the generation shift factor matrix, and

$$\hat{\mathbf{P}}_g^{\max(\min)} = \left[\hat{P}_{g_1}^{\max(\min)}, \ldots, \hat{P}_{g_{N_b}}^{\max(\min)}\right]^T \quad (27)$$

$$\mathbf{F}^{\max(\min)} = \left[F_1^{\max(\min)}, \ldots, F_{N_l}^{\max(\min)}\right]^T. \quad (28)$$

As in [4], unbinding inequality constraints are excluded in our sensitivity analysis. Let us define B_g and B_f as the number of binding constraints associated with generation capacity and line capacity, respectively. Then, the KKT conditions of the Ex-ante problem are written as

$$\begin{aligned}&\text{(i)}\ \frac{\partial C_i\left(P_{g_i}\right)}{\partial P_{g_i}} - \lambda + \sum_{j=1}^{B_g}\tau_j A_{ji} \\ &\qquad + \sum_{l=1}^{B_f}\mu_l S_{li} = 0 \quad \forall i = 1, \ldots, N_b \\ &\text{(ii)}\ \sum_{i=1}^{N_b} P_{g_i} = \sum_{i=1}^{N_b} L_{d_i} \\ &\text{(iii)}\ \sum_{i=1}^{N_b} A_{ji}P_{g_i} = \hat{C}_j \quad \forall j = 1, \ldots, B_g \\ &\text{(iv)}\ \sum_{i=1}^{N_b} S_{li}\left[P_{g_i} - L_{d_i}\right] = D_l \quad \forall l = 1, \ldots, B_f.\end{aligned}$$

after which the above KKT equations are perturbed with respect to P_{g_i}, L_{d_i}, $\hat{C}_j$, λ, τ_j, and μ_j as follows:

$$\begin{aligned}&\text{(i)}\ \underbrace{\frac{\partial}{\partial P_{g_i}}\left(\frac{\partial C_i\left(P_{g_i}\right)}{\partial P_{g_i}}\right)}_{M_i} dP_{g_i} - d\lambda + \sum_{j=1}^{B_g} A_{ji}d\tau_j \\ &\qquad + \sum_{l=1}^{B_f} S_{li}d\mu_l = 0 \quad \forall i = 1, \ldots, N_b \\ &\text{(ii)}\ \sum_{i=1}^{N_b} dP_{g_i} = \sum_{i=1}^{N_b} dL_{d_i} \\ &\text{(iii)}\ \sum_{i=1}^{N_b} A_{ji}dP_{g_i} = d\hat{C}_j \quad \forall j = 1, \ldots, B_g \\ &\text{(iv)}\ \sum_{i=1}^{N_b} S_{li}dP_{g_i} = \sum_{i=1}^{N_b} S_{li}dL_{d_i} \quad \forall l = 1, \ldots, B_f.\end{aligned}$$

It should be noted that the variables D_l, A_{ji}, and S_{li} in the KKT equations are not perturbed. This is due to the fact that: 1) the limits of line flow constraint limits in the Ex-ante model are not updated by the state estimator and 2) the network topology is not

affected by corrupted analog data. These perturbation equations can be expressed in matrix form as

$$\underbrace{\begin{bmatrix} \mathbf{M} & -\mathbf{1}_{N_b} & \Upsilon \\ \mathbf{1}_{N_b}^T & 0 & \mathbf{0} \\ \Upsilon^T & \mathbf{0} & \mathbf{0} \end{bmatrix}}_{\boldsymbol{\Xi}} \begin{bmatrix} d\mathbf{P}_g \\ d\lambda \\ d\boldsymbol{\tau}_s \\ d\boldsymbol{\mu}_s \end{bmatrix} = \underbrace{[\,\mathbf{U_1}^T \quad \mathbf{U_2}^T\,]}_{\boldsymbol{\Phi}} \begin{bmatrix} d\mathbf{L}_d \\ d\hat{\mathbf{C}}_s \end{bmatrix} \tag{29}$$

where

$$\mathbf{M}_{(N_b \times N_b)} = \operatorname{diag}(M_1, \ldots, M_{N_b}) \tag{30}$$

$$\Upsilon_{(N_b \times [B_g + B_f])} = \begin{bmatrix} \mathbf{A}_s^T & \mathbf{B}_s^T \end{bmatrix} \tag{31}$$

$$\mathbf{U_1}_{(N_b \times [N_b+1+B_g+B_f])} = \begin{bmatrix} \mathbf{0} & \mathbf{1}_{N_b} & \mathbf{0} & \mathbf{B}_s^T \end{bmatrix} \tag{32}$$

$$\mathbf{U_2}_{(B_g \times [N_b+1+B_g+B_f])} = \begin{bmatrix} \mathbf{0} & \mathbf{0}_{B_g} & \mathbf{I}_{B_g} & \mathbf{0} \end{bmatrix}. \tag{33}$$

Taking the inverse of $\boldsymbol{\Xi}$ on both sides of (29) yields

$$\begin{bmatrix} d\mathbf{P}_g \\ d\lambda \\ d\boldsymbol{\tau}_s \\ d\boldsymbol{\mu}_s \end{bmatrix} = \underbrace{\boldsymbol{\Xi}^{-1}\boldsymbol{\Phi}}_{\boldsymbol{\Lambda}_p} \begin{bmatrix} d\mathbf{L}_d \\ d\hat{\mathbf{C}}_s \end{bmatrix}. \tag{34}$$

The subscript s of the variables in (29), (31), and (32) represents the subvector (submatrix) of the original vector (matrix) that corresponds to the binding constraints. The matrix $\boldsymbol{\Lambda}_p$ in (34) is partitioned into two sensitivity matrices—$\boldsymbol{\Lambda}_{\mathbf{L}_d}$ and $\boldsymbol{\Lambda}_{\hat{\mathbf{C}}_s}$, shown as

$$\boldsymbol{\Lambda}_p = [\,\boldsymbol{\Lambda}_{\mathbf{L}_d} \;\big|\; \boldsymbol{\Lambda}_{\hat{\mathbf{C}}_s}\,] = \left[\begin{array}{c|c} \frac{\partial \mathbf{P}_g}{\partial \mathbf{L}_d} & \frac{\partial \mathbf{P}_g}{\partial \hat{\mathbf{C}}_s} \\ \frac{\partial \lambda}{\partial \mathbf{L}_d} & \frac{\partial \lambda}{\partial \hat{\mathbf{C}}_s} \\ \frac{\partial \boldsymbol{\tau}_s}{\partial \mathbf{L}_d} & \frac{\partial \boldsymbol{\tau}_s}{\partial \hat{\mathbf{C}}_s} \\ \frac{\partial \boldsymbol{\mu}_s}{\partial \mathbf{L}_d} & \frac{\partial \boldsymbol{\mu}_s}{\partial \hat{\mathbf{C}}_s} \end{array}\right]. \tag{35}$$

Using the sensitivities of two shadow prices with respect to $\hat{\mathbf{C}}_s((\partial\lambda/\partial\hat{\mathbf{C}}_s), (\partial\boldsymbol{\mu}_s/\partial\hat{\mathbf{C}}_s))$ in $\boldsymbol{\Lambda}_{\hat{\mathbf{C}}_s}$ and according to the definition of LMP [14], we finally construct the matrix $\boldsymbol{\Lambda}_A$.

On the other hand, in the Ex-post model, (29) can be extended as follows:

$$\begin{bmatrix} \mathbf{M} & -\mathbf{1}_{N_b} & \Upsilon \\ \mathbf{1}_{N_b}^T & 0 & \mathbf{0} \\ \Upsilon^T & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} d\mathbf{P}_g \\ d\lambda \\ d\boldsymbol{\tau}_s \\ d\boldsymbol{\mu}_s \end{bmatrix} = [\,\mathbf{U_1}^T \quad \mathbf{U_2}^T \quad \mathbf{U_3}^T\,] \begin{bmatrix} d\hat{\mathbf{P}}_g \\ d\hat{\mathbf{C}}_s \\ d\hat{\mathbf{D}}_s \end{bmatrix}. \tag{36}$$

$\hat{\mathbf{D}}_s$ is the subvector of $\hat{\mathbf{D}}$ (the real power flow estimate vector) that corresponds to the binding constraints, and

$$\mathbf{U_3}_{(B_f \times [N_b+1+B_g+B_f])} = \begin{bmatrix} \mathbf{0} & \mathbf{0}_{B_f} & \mathbf{0} & \mathbf{I}_{B_f} \end{bmatrix}. \tag{37}$$

Compared with (35), $\boldsymbol{\Lambda}_p$ in the Ex-post model is written as

$$\boldsymbol{\Lambda}_p = [\,\boldsymbol{\Lambda}_{\hat{\mathbf{P}}_g} \;\big|\; \boldsymbol{\Lambda}_{\hat{\mathbf{C}}_s} \;\big|\; \boldsymbol{\Lambda}_{\hat{\mathbf{D}}_s}\,] = \left[\begin{array}{c|c|c} \frac{\partial \mathbf{P}_g}{\partial \hat{\mathbf{P}}_g} & \frac{\partial \mathbf{P}_g}{\partial \hat{\mathbf{C}}_s} & \frac{\partial \mathbf{P}_g}{\partial \hat{\mathbf{D}}_s} \\ \frac{\partial \lambda}{\partial \hat{\mathbf{P}}_g} & \frac{\partial \lambda}{\partial \hat{\mathbf{C}}_s} & \frac{\partial \lambda}{\partial \hat{\mathbf{D}}_s} \\ \frac{\partial \boldsymbol{\tau}}{\partial \hat{\mathbf{P}}_g} & \frac{\partial \boldsymbol{\tau}}{\partial \hat{\mathbf{C}}_s} & \frac{\partial \boldsymbol{\tau}}{\partial \hat{\mathbf{D}}_s} \\ \frac{\partial \boldsymbol{\mu}}{\partial \hat{\mathbf{P}}_g} & \frac{\partial \boldsymbol{\mu}}{\partial \hat{\mathbf{C}}_s} & \frac{\partial \boldsymbol{\mu}}{\partial \hat{\mathbf{D}}_s} \end{array}\right]. \tag{38}$$

B. Sensitivity of State Estimation to SCADA Data

Sensitivity analysis of state estimation subject to SCADA measurements was pioneered by Stuart and Herget [15], who investigated the effect of power system modeling errors on weighted least-squares (WLS) state estimation. A more rigorous sensitivity analysis method, based on the same perturbation approach illustrated in [4], has been proposed by Mínguez and Conejo [16]. This method has been formulated in a general optimization problem that allows for the sensitivity analysis of alternative state estimation methods with different objective functions, such as the least absolute value (LAV) from the WLS. It should be noted that, in this paper, the sensitivity analysis is based on WLS state estimation. However, one can apply it to various state estimation methods by using the method proposed in [16].

Here, we first derive the matrix $\boldsymbol{\Lambda}_B$ that illustrates the sensitivities of the real power injection and real flow measurement estimates with respect to the changes in all types of measurements. In (4), the matrix $\boldsymbol{\Psi}(\hat{\mathbf{x}}^k)$ is defined and partitioned as

$$\boldsymbol{\Psi}(\hat{\mathbf{x}}^k) = \left[\mathbf{G}(\hat{\mathbf{x}}^k)\right]^{-1} \mathbf{H}^T(\hat{\mathbf{x}}^k)\mathbf{R}^{-1} = \left[\frac{\boldsymbol{\Psi}_{\hat{\boldsymbol{\theta}}}(\hat{\mathbf{x}}^k)}{\boldsymbol{\Psi}_{\hat{\mathbf{V}}}(\hat{\mathbf{x}}^k)}\right] \tag{39}$$

where $\boldsymbol{\Psi}_{\hat{\boldsymbol{\theta}}}(\hat{\mathbf{x}}^k)$ and $\boldsymbol{\Psi}_{\hat{\mathbf{V}}}(\hat{\mathbf{x}}^k)$ represent the sensitivities of the voltage phase angle estimates and the magnitudes with respect to all perturbed measurements at the k-th iteration, respectively. Therefore, (4) can be rewritten as

$$\begin{bmatrix} d\hat{\boldsymbol{\theta}}^{k+1} \\ d\hat{\mathbf{V}}^{k+1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Psi}_{\hat{\boldsymbol{\theta}}}(\hat{\mathbf{x}}^k) \\ \boldsymbol{\Psi}_{\hat{\mathbf{V}}}(\hat{\mathbf{x}}^k) \end{bmatrix} d\mathbf{z}. \tag{40}$$

It should be noted that the DCOPF-based SCED is formulated with linearized real power injection and a flow estimation solution [17]. Using the linear equations in the upper partition of (40) and the matrix $\boldsymbol{\Psi}_{\hat{\boldsymbol{\theta}}}$ computed with the converged estimate $\hat{\mathbf{x}}$, we have the following sensitivity equation:

$$d\hat{\mathbf{z}}_r = \begin{bmatrix} \mathbf{B}_{P\theta}^S \\ \mathbf{B}_{P\theta} \\ \mathbf{B}_{F\theta} \end{bmatrix} d\hat{\boldsymbol{\theta}} = \begin{bmatrix} \mathbf{B}_{P\theta}^S \\ \mathbf{B}_{P\theta} \\ \mathbf{B}_{F\theta} \end{bmatrix} \boldsymbol{\Psi}_{\hat{\boldsymbol{\theta}}} d\mathbf{z} = \mathbf{K} d\mathbf{z} \tag{41}$$

where

$$\mathbf{K} = \begin{bmatrix} \mathbf{B}_{P\theta}^S \\ \mathbf{B}_{P\theta} \\ \mathbf{B}_{F\theta} \end{bmatrix} \boldsymbol{\Psi}_{\hat{\boldsymbol{\theta}}}. \tag{42}$$

$d\hat{\mathbf{z}}_r$ is the perturbed estimate vector of the real power injection and the flow measurements. The matrix $\mathbf{B}_{P\theta} = \mathbf{A}_r\mathbf{B}_d\mathbf{A}_r^T$ is defined as the $(N_b - 1) \times (N_b - 1)$ reduced node-to-node susceptance matrix that explains the relationship between real power injections at any bus except the slack bus and the phase angles. Here, $\mathbf{B}_d = \operatorname{diag}(s_1, s_2, \ldots, s_{N_l})$ is the $N_l \times N_l$ diagonal branch susceptance matrix and $\mathbf{A}_r$ is the $(N_b - 1) \times N_l$ reduced node-to-branch incidence matrix without a slack bus. According to the law of conservation of power, the $1 \times (N_b - 1)$ matrix $\mathbf{B}_{P\theta}^S = -\mathbf{1}_{(N_b-1)}^T\mathbf{B}_{P\theta}$ is derived, and it explains the relationship between real power injections at the slack bus and the

TABLE I
GENERATOR PARAMETERS IN THE IEEE 14-BUS SYSTEM

Bus	$P_{g_i}^{\min}$(MW)	$P_{g_i}^{\max}$(MW)	a_i(\$/MWh)	b_i(\$/(MW)^{2}h)
1	0	332.4	20	0.043
2	0	140	20	0.25
3	0	100	40	0.01
6	0	100	40	0.01
8	0	100	40	0.01

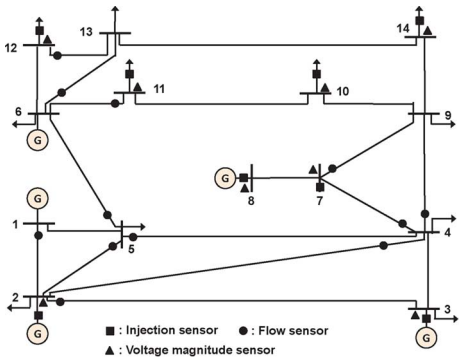

Fig. 2. IEEE 14-bus system with a given measurement configuration.

phase angles. The matrix $\mathbf{B}_{F\theta} = \mathbf{B}_d\mathbf{A}_r^T$ specifies the relationship between real power flows and the phase angles. Using (42), we compute the matrix $\mathbf{\Lambda}_B = \mathbf{K}$.

Remark 1: The proposed sensitivity analysis is beneficial in several ways. Through online predictions of LMP variations, system operators can identify economically sensitive buses with respect to data corruption. This could help system operators prioritize sensor data quality upgrades in view of robust market operations. From a practical perspective, the proposed approach can be easily integrated into applications in existing EMSs or market management systems (MMSs).

Remark 2: The sensitivity index (20) can be used to provide the following *system-wide* metrics that alert system operators to the jth most and kth least influential sensors on LMP, on average, with respect to a total of N_b buses and T dispatch intervals:

$$j = \arg\max_j \left(\sum_{i=1}^{N_b} \sum_{t=1}^{T} \left| \mathbf{\Lambda}_{(i,j)}^{(t)} \right| / N_b T \right)$$

$$k = \arg\min_k \left(\sum_{i=1}^{N_b} \sum_{t=1}^{T} \left| \mathbf{\Lambda}_{(i,k)}^{(t)} \right| / N_b T \right).$$

V. NUMERICAL EXAMPLE

Here, we illustrate and verify the proposed approach to quantifying the sensitivities of real-time LMP with respect to changes in sensor data. The proposed sensitivity analysis is applied to IEEE 14-bus and 118-bus systems. System data for the IEEE 14-bus system are taken from the MATPOWER 4.0 IEEE 14-bus test case file. Table I shows the generator parameters in the IEEE 14-bus system.

In this simulation, the measurement configuration consists of eight voltage magnitude measurements, eight pairs of real and reactive power injection measurements, and 12 pairs of real and reactive power flow measurements. V_i is the measurement of voltage magnitude at bus i, P_i and Q_i are the measurements of real and reactive power injection at bus i, respectively, and $P_{i,j}$ and $Q_{i,j}$ are the measurements of real and reactive power flow from bus i to bus j, respectively. Fig. 2 shows the IEEE 14-bus system with a measurement configuration that consists of the following five measurement sets:

$$\begin{aligned}
\mathcal{S}_v &= \{V_2, V_3, V_7, V_8, V_{10}, V_{11}, V_{12}, V_{14}\} \\
\mathcal{S}_{ri} &= \{P_2, P_3, P_7, P_8, P_{10}, P_{11}, P_{12}, P_{14}\} \\
\mathcal{S}_{ai} &= \{Q_2, Q_3, Q_7, Q_8, Q_{10}, Q_{11}, Q_{12}, Q_{14}\} \\
\mathcal{S}_{rf} &= \{P_{1,2}, P_{2,3}, P_{4,2}, P_{4,7}, P_{4,9}, P_{5,2}, P_{5,4}, P_{5,6}, \\
&\qquad P_{6,13}, P_{7,9}, P_{11,6}, P_{12,13}\} \\
\mathcal{S}_{af} &= \{Q_{1,2}, Q_{2,3}, Q_{4,2}, Q_{4,7}, Q_{4,9}, Q_{5,2}, Q_{5,4}, Q_{5,6}, \\
&\qquad Q_{6,13}, Q_{7,9}, Q_{11,6}, Q_{12,13}\}.
\end{aligned}$$

In this measurement configuration, the locations of the voltage magnitude measurements are consistent with those of the real and reactive power injection measurements. For each measurement set, the measurement index is numbered from one to the total number of measurements in each set.

We assume that all measurements are corrupted by additive Gaussian noises with equal variances $\sigma^2 = 0.00001$. Finally, for all buses i, j, and k, we compute LMP sensitivities with respect to the five types of measurements—real/reactive power injection, real/reactive power flow, and voltage magnitude—as follows:

$$\frac{\partial \pi_i}{\partial P_j}, \frac{\partial \pi_i}{\partial Q_j}, \frac{\partial \pi_i}{\partial P_{j,k}}, \frac{\partial \pi_i}{\partial Q_{j,k}}, \frac{\partial \pi_i}{\partial V_j}. \tag{43}$$

Units for the sensitivities $\{(\partial\pi_i/\partial P_j), (\partial\pi_i/\partial P_{j,k})\}$, $\{(\partial\pi_i/\partial Q_j), (\partial\pi_i/\partial Q_{j,k})\}$, and $\{(\partial\pi_i/\partial V_j)\}$ are (\$/MWh)/(puMW), (\$/MWh)/(puMVAr), and (\$/MWh)/(puV), respectively.

Fig. 3 provides snapshots of five different Ex-ante LMP sensitivities in (43) at some buses in the IEEE 14-bus system with line 3–4 congestion. These figures provide information about the directions of the post-corruption LMPs as well as their sensitivities with respect to each type of measurement at a given dispatch time. In this simulation, after the Ex-ante dispatch problem has been solved, there exist two binding generation capacity constraints: P_{g_3} and P_{g_8} are binding at $\hat{P}_{g_3}^{\min}$ and $\hat{P}_{g_8}^{\max}$, respectively. We assume that the corruption of the measurements impacts the binding constraint associated with P_{g_3}. In other words, the corrupted measurements affect $\hat{P}_{g_3}^{\min}$ (an intermediate variable in (21)), subsequently leading to changes in all the LMPs. We randomly choose seven buses (buses 1, 2, 3, 4, 5, 10, 13) out of the 14 to differentiate clearly the LMP sensitivities among the various buses. The absolute values of the LMP sensitivities at buses 3 and 5 are the largest and smallest, setting the upper and lower bounds for sensitivity at the fourteen buses. We obtain from the simulation results the following observations.

- (O1) *Sensitivity grouping property*: all buses can be categorized into two sensitivity groups. In each group, buses obtain sensitivities with the same sign, but of different magnitude and subject to all types of measurements. Group I includes buses 1, 2, 3, and 5, and Group II buses 4, 10, and 13.

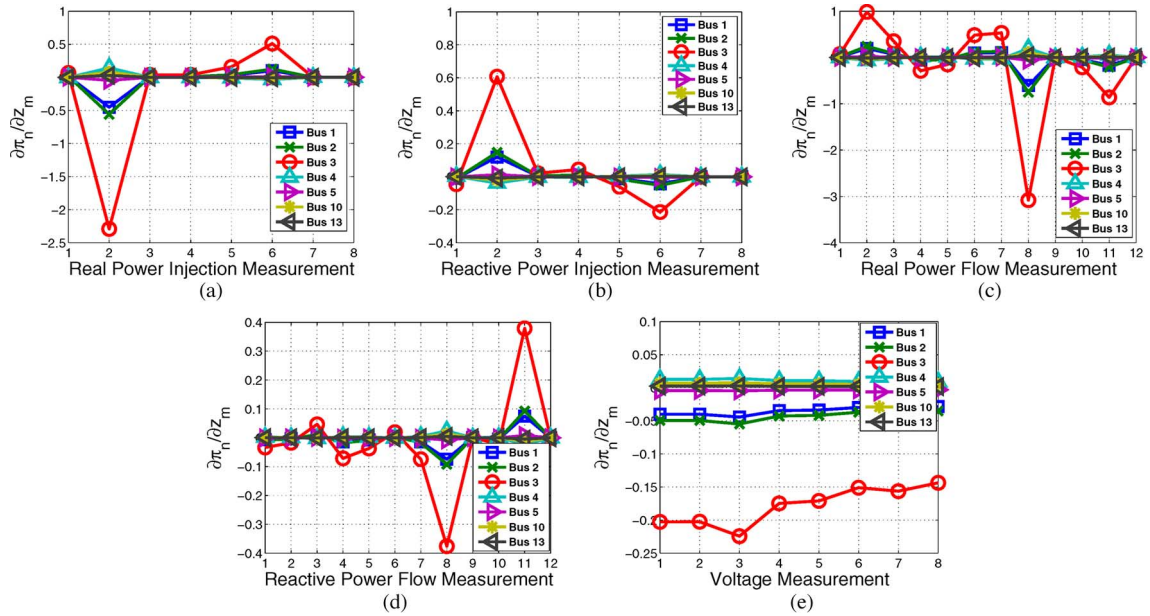

Fig. 3. Sensitivities of Ex-ante prices with respect to (a) real power injection measurements, (b) reactive power injection measurements, (c) real power flow measurements, (d) reactive power flow measurements, and (e) voltage magnitude measurements. Line 3–4 is congested and P_{g_3} is binding at $\hat{P}_{g_3}^{\min}$ in the IEEE 14-bus system.

For example, in Fig. 3(b), the corruption of z_2 yields positive sensitivities for Group I and negative sensitivities for Group II, whereas the corruption of z_6 yields the reverse: negative sensitivities for Group I and positive sensitivities for Group II. This grouping property enables system operators to predict rapidly the direction of LMP's distortion in response to sensor data corruption.

- (O2) *Identification of buses that are economically sensitive to data corruption*: buses incident to both ends of the congested line have the highest LMP sensitivities with respect to sensor data corruption. For example, bus 3 in Group I and bus 4 in Group II incident to congested line 3–4 have the largest absolute sensitivities in each group. In particular, it should be noted that the largest sensitivities are associated with bus 3. This implies that bus 3 is the most financially vulnerable to any corruption in sensor measurement.
- (O3) *Identification of influential sensors on LMP*: the sensor most influential on LMP change is identified in each measurement group. In Fig. 3(a) and (b), the sensors with z_2 (P_3 and Q_3) have the most significant impact on LMP. This is due to the fact that the change of the intermediate variable $\hat{P}_{g_3}^{\min}$ is dominantly affected by P_3 and Q_3, subsequently leading to more change in LMP. This effect is also verified in Fig. 9(a) and (b) based on the IEEE-118 bus system. In Fig. 3(c)–(e), the sensors with z_8, z_{11} and z_3 ($P_{5,6}$, $Q_{11,6}$ and V_7) are the most influential, respectively. In addition, it should be noted that the localized effects on increasing sensitivity of measurements adjacent to the congested line and/or the intermediate variable do not always hold true. For example, $z_{11}(P_{11,6})$ is farther away from both the congested line and the intermediate variable than $z_5(P_{4,9})$; however, in Fig. 3(c), data corruption in the former leads to a higher sensitivity than in the latter. This nonlocalized data effect motivates system operators to use our developed tool for identifying which sensors impact LMP sensitivity.
- (O4) *Impact of different types of sensor data on LMP*: through a comparison of all of the figures, LMP appears to be more sensitive to real power injection/flow measurements than to reactive power injection/flow and voltage magnitude measurements. In order to compare the sensitivities of different units fairly, a normalized LMP sensitivity $|z_j|(\partial\pi_i/\partial z_j)$ is defined, which is incorporated into the following proposed metric:

$$\Omega_k^i = \sum_{j=1}^{|\mathcal{S}_k|} \left| |z_j| \frac{\partial \pi_i}{\partial z_j} \right| / |\mathcal{S}_k| \tag{43}$$

where Ω_k^i is the average of the absolute normalized sensitivities at bus i with respect to any measurement z_j in the set $\mathcal{S}_k (k = v, ri, ai, rf, af)$. The cardinality of the set $|\mathcal{S}_k|$ means the number of elements in $\mathcal{S}_k$. For example, at bus 3, we compute $\Omega_{ri}^3 = 0.474$, $\Omega_{rf}^3 = 0.253$, $\Omega_v^3 = 0.175$, $\Omega_{af}^3 = 0.013$, and $\Omega_{ai}^3 = 0.012$, which is consistent with our expectation that real power injection and flow measurements have a more significant impact on LMP sensitivity than other measurements. This is due to the fact that DCOPF-based SCED is conducted based on a linearized state estimation solution that is more influenced by real power measurements than by reactive power and voltage magnitude measurements, as illustrated in (40) and (41).

- (O5) In Fig. 3(e), LMP sensitivities at all buses affected by corrupted voltage magnitude measurements fluctuate more smoothly than the ones affected by other types of corrupted measurements. In other words, all voltage magnitude measurements impact LMP variations almost evenly.

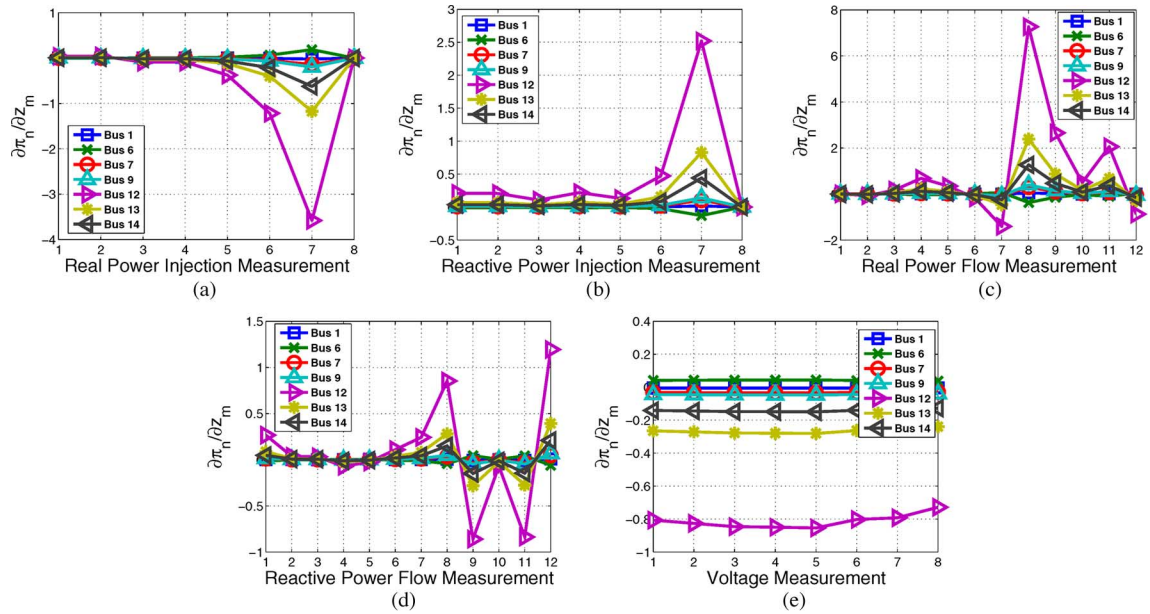

Fig. 4. Sensitivities of Ex-post prices with respect to (a) real power injection measurements, (b) reactive power injection measurements, (c) real power flow measurements, (d) reactive power flow measurements, and (e) voltage magnitude measurements. Line 6–12 is congested and the corresponding line flow is binding at the capacity limit of line 6–12 in the IEEE 14-bus system.

In addition, the nonlocalized effect mentioned in (O3) is also verified between $z_2(V_3)$ and $z_3(V_7)$.

Fig. 4 provides snapshots of the Ex-post LMP sensitivities at arbitrarily chosen buses (buses 1, 6, 7, 9, 12, 13, and 14) with respect to the aforementioned five types of sensor measurements. In this simulation, line 6–12 is assumed to be congested at both Ex-ante dispatch and Ex-post dispatch. $\hat{P}_{6,12}$ is chosen as an intermediate variable to compute LMP sensitivity. We can observe from Fig. 4 the same phenomena as in Fig. 3: (O1) Group 1 for buses 1 and 6, and Group 2 for buses 7, 9, 12, 13, and 14; (O2) buses 6 and 12 incident to the congested line have the largest absolute value of LMP sensitivity in each group; (O3) in Fig. 4(a) and (b), the sensors with z_7 (P_{12} and Q_{12}) have the most significant impact on LMP, and in Fig. 4(c)–(e), the sensors with z_8, z_{12}, and z_5 ($P_{5,6}$, $Q_{12,13}$, and V_{10}) are the most influential, respectively; and (O4) and (O5) real power measurements have a stronger impact on LMP sensitivity than the reactive power and voltage magnitude measurements, and the voltage magnitude measurements influence LMP sensitivity almost evenly. In addition, these analytical LMP sensitivity results have been checked using the perturbation method ($f'(x) \approx ([f(x+\epsilon) - f(x)]/\epsilon)$). It has been verified that the sensitivity results obtained from the proposed analytical method are consistent with those from the perturbation method.

Fig. 5 shows actual Ex-ante LMP and how they differ when they have or do not have corrupted data at all buses. It is assumed that the magnitude of z_8 is corrupted by 2% in Fig. 3(c). In the Chi-squares test [12] within a 99% confidence level, the estimated objective functions and the bad data detection threshold are computed. $J(\hat{x}) = 15.69$ and $J^{(b)}(\hat{x}) = 30.17$ correspond to the values of the estimated objective functions without and with corrupted data, respectively, and $\chi^2 = 38.93$ is the value of the bad data detection threshold. It should be noted that, since

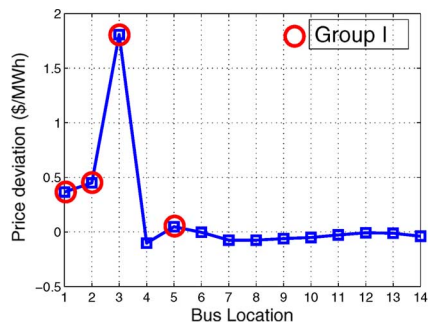

Fig. 5. LMP differences between with and without corrupted data when z_8 is corrupted in Fig. 3(c).

$J^{(b)}(\hat{x}) = 30.17 < \chi^2 = 38.93$, the corrupted measurement z_8 bypasses the bad data detection engine, which could then lead to LMP distortion. As expected, Fig. 5 justifies the result of our sensitivity analysis in two main ways. First, the prices at buses 1, 2, 3, and 5 in Group I change in a positive direction whereas the prices at the buses in Group II change in a negative direction. This observation explains the grouping property specified in (O1). Second, the descending order of the magnitudes of the actual price deviations is in accordance with that of sensitivity magnitudes. For example, Fig. 3(c) shows that buses 3, 2, 1, and 5 in Group I are in descending order of sensitivity magnitudes, which is consistent with the descending order of the actual price deviations at those same buses in Fig. 5.

Fig. 6 shows the Ex-ante LMP deviations that are caused by the undetectable same amount of corruption in each measurement group $\{P_3, Q_3, V_3\}$ and $\{P_{5,6}, Q_{5,6}\}$. These figures show that real power injection and flow measurements have a more significant impact on LMP than other measurements. This fact justifies observation (O4).

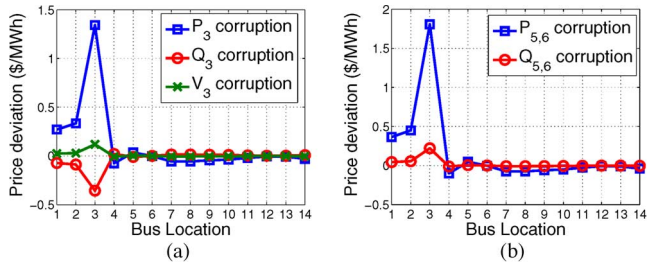

Fig. 6. LMP differences between with and without corrupted data in Fig. 3. (a) P_3, Q_3, and V_3 corruptions. (b) $P_{5,6}$ and $Q_{5,6}$ corruptions.

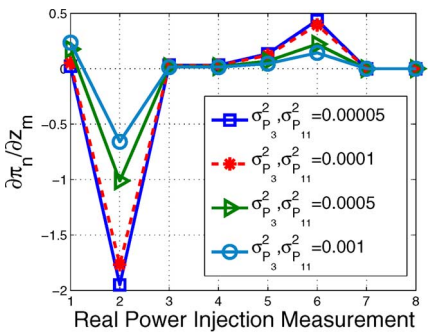

Fig. 7. Comparison of LMP sensitivities at bus 3 in Fig. 3(a) with varying variances of injection measurements P_3 and P_{11}.

Fig. 7 shows the impact of sensor measurement accuracy on LMP sensitivity. In this figure, four plots represent LMP sensitivities at bus 3 in Fig. 3(a), with consistently varying variances of the two injection measurements $z_2(P_3)$ and $z_6(P_{11})$. These sensitivities are measured at four different variance levels; $\sigma^2 = 0.00005, 0.0001, 0.0005$, and 0.001. We can observe from Fig. 7 that, as the measurement variance decreases (i.e., the measurement accuracy increases), the corresponding LMP sensitivity increases. In other words, more accurate sensors lead to more change in LMP while sensor data remain corrupted. This shows the coupling between state estimation accuracy and LMP calculation. Based on this observation, one possible guideline for mitigating the financial risk from data corruption is to make it a high priority to protect accurate sensors.

For the IEEE 118-bus system, with 54 generation buses and 186 transmission lines as shown in Fig. 8, we assume that real and reactive power injection measurements are placed at 49 generator buses, voltage magnitude measurements at nine generator buses, and real and reactive flow measurements at 129 lines. Therefore, this system has a total of 365 measurements. System data for the IEEE 118-bus system are taken from the MATPOWER 4.0 IEEE 118-bus test case file.

Fig. 9 show the Ex-ante LMP sensitivities at buses 15, 17, 35, and 75 in the IEEE 118-bus system with line 15–17 congestion with respect to the five different types of measurement. The magnitudes of the sensitivities at buses 15 and 17 are the highest in each sensitivity group. $\hat{P}_{19}^{\max}$ is chosen as an intermediate variable to compute LMP sensitivity. As expected, all observations from Fig. 3 are also verified in the larger IEEE 118-bus system: 1) sensor grouping property (Group 1: buses 15 and 35, and Group 2: buses 17 and 75); 2) identification of the most economically sensitive buses in each group (buses 15 and 17) and the most influential sensors (e.g., z_{10} with P_{19} and Q_{19} in Fig. 9(a) and (b)) on LMP change; and 3) the impact of different types of sensor data on LMP (e.g., the more significant impact of real power measurements than other types of measurements).

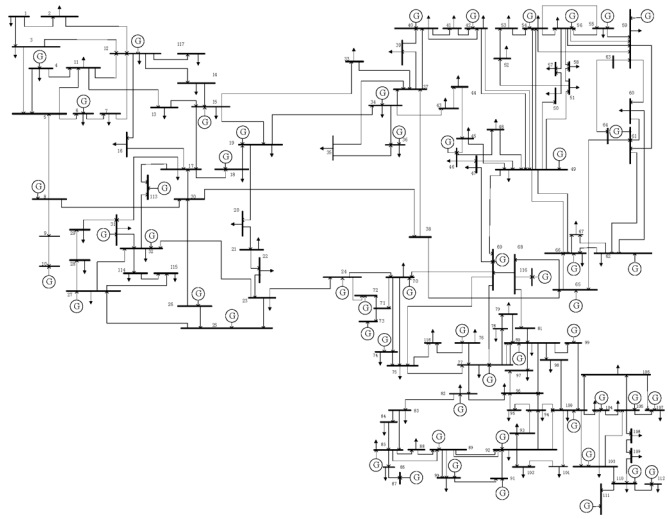

Fig. 8. IEEE 118-bus system.

VI. Discussions

A. Phasor Measurement Unit (PMU) Implementation

Recently, with PMUs being increasingly deployed in power systems, novel hybrid state estimation methods based on traditional SCADA and PMU measurements have been intensively investigated [18]–[21]. The state estimation measurement model (1) presented in this paper can be easily expanded into a hybrid model (e.g., [19, eq. (12)]). Using the same steps illustrated in Section IV-B, our proposed approach quantifies LMP sensitivity with respect to PMU as well as SCADA measurements.

B. Dependent Sensor Measurements

Recent studies have speculated that state estimation measurements with a substation may be correlated and proposed as a response a novel dependent WLS (DWLS) state estimation method that considers measurement dependencies [22]. These methods differ from the traditional WLS method based on independent measurements only in that it computes the nondiagonal covariance matrix of dependent measurements using the point estimate technique in the WLS formulation. By simply replacing diagonal covariance matrix $\mathbf{R}$ in (2) with a computed nondiagonal covariance matrix, our proposed formulation is also applicable to any LMP sensitivity analysis subject to dependent sensor measurements.

C. LMP Sensitivity With Respect to Network Topology Errors

This paper is limited to the study of LMP sensitivity with respect to changes in the analog data corruption-induced power flow estimate. A very important extension of our work here

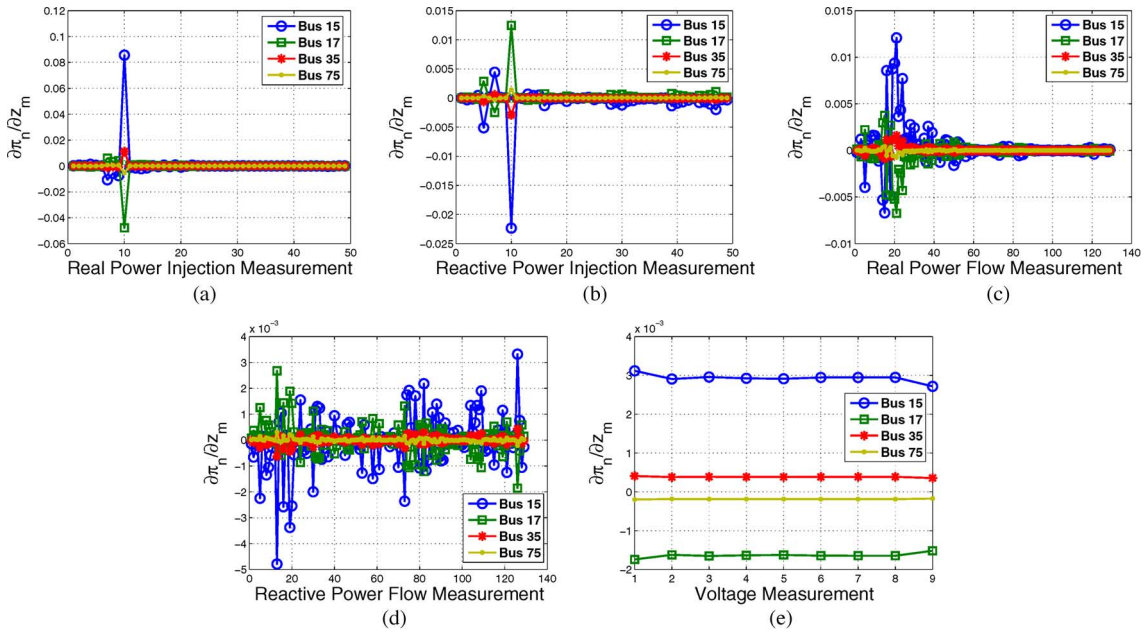

Fig. 9. Sensitivities of Ex-ante prices with respect to (a) real power injection measurements, (b) reactive power injection measurements, (c) real power flow measurements, (d) reactive power flow measurements, and (e) voltage magnitude measurements. Line 15–17 is congested and $P_{g_{19}}$ is binding at $\bar{P}_{g_{19}}^{\max}$ in the IEEE 118-bus system.

would be to study LMP sensitivity with respect to network topology estimate changes caused by corrupted digital sensor data such as the on/off status of a circuit breaker. A key part of this task would be to investigate how such LMP sensitivity is analytically expressed as a function of topology information such as the generation shift factor matrix and corrupted digital data. As an initial step in this direction, recent work of ours [23] proposed a simple LMP sensitivity index that accounts for the relationship between changes in network topology and LMP during single transmission line congestion. This index quantifies the *economic* impact of network topology errors on LMP and is expressed in terms of the energy costs of marginal units and the congested line-related generation shift factors at any bus and marginal units. Here, a marginal unit is defined as a unit that generates power in a range somewhere between its minimum and maximum capacity. Future work should include the development of a sensitivity index that evaluates the *cyber* impact of digital data on network topology estimate.

VII. Conclusion

In this paper, we present an analytical framework for calculating LMP sensitivity in response to small variations in SCADA and/or PMU measurement data. Corrupted sensor data are shown to deviate power system state estimation from their actual values, which subsequently leads to the distortion of real-time market LMPs. We build two matrices: the first with LMP sensitivity at any bus to any estimate, and the second with sensitivity of any estimate to data at any sensor. A unified matrix that combines these two matrices in multiplication form enables system operators to quantify the impact on LMP of data at any sensor at any bus throughout the entire transmission network. Our simulation results suggest that the proposed sensitivity matrix can provide system operators with a quick and accurate method to identify the buses most vulnerable to measurement errors. In addition, we verify that more accurate sensors impact LMP much more significantly.

References

[1] F. C. Schweppe, J. Wildes, and D. B. Rom, "Power system static state estimation, Parts I, II and III," *IEEE Trans. Power App. Syst.*, vol. PAS-89, no. 1, pp. 120–135, Jan. 1970.

[2] W. W. Hogan, "Contract networks for electric power transmission," *J. Regulatory Econ.*, vol. 4, no. 3, pp. 211–242, Sep. 1992.

[3] T. Zheng and E. Litvinov, "On ex post pricing in the real-time electricity market," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 153–164, Feb. 2011.

[4] A. J. Conejo, E. Castillo, R. Mínguez, and F. Milano, "Locational marginal price sensitivities," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 2026–2033, Nov. 2005.

[5] F. Li and R. Bo, "DCOPF-based LMP simulation: Algorithm, comparison with ACOPF, and sensitivity," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 1475–1485, Nov. 2007.

[6] F. Li, "Continuous locational marginal pricing (CLMP)," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 1638–1646, Nov. 2007.

[7] R. Bo and F. Li, "Probabilistic LMP forecasting considering load uncertainty," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1279–1289, Aug. 2009.

[8] Y. Liu, M. K. Reiter, and P. Ning, "False data injection attacks against state estimation in electric power grids," in *Proc. 16th ACM Conf. Comput. Commun. Security*, Nov. 2009, pp. 21–32.

[9] O. Kosut, L. Jia, R.-J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Trans. Smart Grids*, vol. 2, no. 4, pp. 645–658, Dec. 2011.

[10] 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.

[11] L. Jia, R. J. Thomas, and L. Tong, "Malicious data attack on real-time electricity market," in *Proc. Int. Conf. Acoust., Speech Signal Process.*, May 2011, pp. 5952–5955.

[12] A. Abur and A. G. Expósito, *Power System State Estimation. Theory and Implementation*. New York, NY, USA: Marcel Dekker, 2004.

[13] F. Li, Y. Wei, and S. Adhikari, "Improving an unjustified common practice in Ex Post LMP Calculation," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 1195–1197, May 2010.

[14] F. F. Wu, P. Varaiya, P. Spiller, and S. Oren, "Folk theorems on transmission access: Proofs and counterexamples," *J. Regulatory Econ.*, vol. 10, no. 1, pp. 5–23, Jul. 1996.

[15] T. A. Stuart and C. J. Herget, "A sensitivity analysis of weighted least squares state estimation for power systems," *IEEE Trans. Power App Syst.*, vol. PAS-92, no. 5, pp. 1696–1701, Sep. 1973.
[16] R. Mínguez and A. J. Conejo, "State estimation sensitivity analysis," *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 1080–1091, Aug. 2007.
[17] 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.
[18] 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.
[19] T. S. Bi, X. H. Qin, and Q. X. Yang, "A novel hybrid state estimator for including synchronized phasor measurements," *Electr. Power Syst. Res.*, vol. 81, no. 5, pp. 1343–1352, Feb. 2008.
[20] G. Valverde, S. Chakrabarti, E. Kyriakides, and V. Terzija, "A constrained formulation for hybrid state estimation," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1102–1109, Aug. 2011.
[21] K. Das, J. Hazra, D. P. Seetharam, R. K. Reddi, and A. K. Sinha, "Real-time hybrid state estimation incorporating SCADA and PMU measurements," in *Proc. 3rd IEEE PES Int. Conf. Exhibition on Innovative Smart Grid Technol.*, Berlin, Germany, Oct. 2012, pp. 1–8.
[22] E. Caro, A. J. Conejo, and R. Mínguez, "Power system state estimation considering measurement dependencies," *IEEE Trans. Power Syst.*, vol. 24, no. 4, pp. 1875–1885, Nov. 2009.
[23] D.-H. Choi and L. Xie, "Impact analysis of locational marginal price subject to power system topology errors," in *Proc. 4th IEEE Int. Conf. Smart Grid Commun.*, Vancouver, BC, Canada, to be published.

Dae-Hyun Choi (S'10) received the B.S. degree in electrical engineering from Korea University, Seoul, Korea, in 2002, and the M.Sc. degree in electrical and computer engineering from Texas A&M University, College Station, TX, USA, in 2008, where he is currently working toward the Ph.D. degree in the Department of Electrical and Computer Engineering.

From 2002 to 2006, he was a Researcher with Korea Telecom (KT), Seoul, Korea, where he worked on designing and implementing home network systems. His research interests include power system state estimation, electricity markets, the cyber-physical security of smart grids, and the theory and application of cyber-physical energy systems.

Le Xie (S'05–M'10) received the B.E. degree in electrical engineering from Tsinghua University, Beijing, China, in 2004, the M.Sc. degree in engineering sciences from Harvard University, Cambridge, MA, USA, in 2005, and the Ph.D. degree from Carnegie Mellon University, Pittsburgh, PA, USA, in 2009.

He is an Assistant Professor with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA, where he is affiliated with the Electric Power and Power Electronics Group. His industry experience includes an internship in 2006 at ISO-New England and an internship in 2007 at Edison Mission Energy Marketing and Trading. His research interests include the modeling, estimation, and control of large-scale power systems, and electricity markets.