Covariance Analysis of LAV Robust Dynamic State Estimation in Power Systems

Lu Sun, Tengpeng Chen, Weng Khuen Ho, Keck Voon Ling and Jan M. Maciejowski, *Fellow*, *IEEE*

Abstract—In power system state estimation, the robust Least Absolute Value robust dynamic estimator is well-known. However, the covariance of the state estimation error cannot be obtained easily. In this paper, an analytical equation is derived using Influence Function approximation to analyze the covariance of the robust Least Absolute Value dynamic state estimator. The equation gives insights into the precision of the estimation and can be used to express the variances of the state estimates as functions of measurement noise variances, enabling the selection of sensors for specified estimator precision. Simulations on the IEEE 14-bus, 30-bus and 118-bus systems are given to illustrate the usefulness of the equation. Monte-Carlo experiments can also be used to determine the covariance, but many data points are needed and hence many runs are required to achieve convergence. Our result shows that to obtain the covariance of the state estimation error, the analytical equation proposed in this paper is four-order of magnitude faster than a 10,000-run Monte-Carlo experiment on both the IEEE 14-bus and 30-bus systems.

Index Terms—Phasor Measurement Unit; Dynamic State Estimation; Innovation Model; Least Absolute Value; Influence Function.

NOMENCLATURE

k — Time index
$x(k)$ — True voltage phasors, dimension $n \times 1$
F — Process matrix, dimension $n \times n$
G — Input matrix
$u(k)$ — Control signal
$z(k)$ — Measurement vector, dimension $m \times 1$
H — Output matrix, dimension $m \times n$
$w(k)$ — Process noise at k, dimension $n \times 1$
Q — Process noise covariance matrix, dimension $n \times n$
$v(k)$ — Measurement noise at k, dimension $m \times 1$
R — Measurement noise covariance, dimension $m \times m$

Lu Sun and W.K. Ho are with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117576. (e-mail: sunlu@u.nus.edu, wk.ho@nus.edu.sg).
Tengpeng Chen is with the Department of Instrumental and Electrical Engineering, Xiamen University, China 361102. (e-mail: tpchen@xmu.edu.cn).
K.V. Ling is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798. (e-mail: ekvling@ntu.edu.sg).
Jan M. Maciejowski is with the Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, United Kingdom. (e-mail: jmm@eng.cam.ac.uk).

$x'(k)$ — Innovation model state vector, dimension $n \times 1$
Γ — Innovation model input matrix, dimension $n \times m$
$\epsilon(k)$ — Innovation vector, dimension $m \times 1$
σ_i — Standard deviation of $\epsilon_i(k)$
ρ_{ij} — Correlation of $\epsilon_i(k)$ and $\epsilon_j(k)$
ϵ — $\epsilon = [\epsilon(1)^T, ..., \epsilon(N)^T]^T$
K_∞ — Steady-state Kalman gain
J — Cost function
$e_i(k)$ — Measurement residual
E — $E = [e(1)^T, ..., e(N)^T]^T$
N — Measurement batch size
$\hat{x}'(k)$ — The state estimates at k
Ψ — Partial differential of J over $\hat{x}'(N)$
$\text{IF}(\cdot)$ — Influence Function
$\tilde{x}'(k)$ — Innovation model estimation error $\hat{x}'(k) - x'(k)$
$\tilde{x}(k)$ — Estimation error $\hat{x}'(k) - x(k)$
$\Delta x(k)$ — Innovation model error $x'(k) - x(k)$
P — Covariance of $\Delta x(k)$
$f(\epsilon)$ — Joint Probability Density Function of ϵ
$\gamma(\cdot)$ — Cost function of LAV estimator

I. INTRODUCTION

A. Motivation

Traditionally power systems are assumed to be at steady-state and for simplicity state estimation was designed for a static system [1]. However, under normal circumstances, the states change slowly and continuously [2]. To capture their dynamic behaviours and monitor the system, dynamic state estimation techniques have been developed [3], [4]. Dynamic state estimators have the dual advantages of being more accurate and the ability to predict the state of the system one-step ahead [5]. In the past, Power System State Estimation (PSSE) was treated as a nonlinear problem. In recent years, with the usage of Phasor Measurement Unit (PMU) which measures the phasor of the bus voltages and line currents directly, it is transformed into a linear problem [6].

The Gaussian noise assumption is commonly made in PSSE problems [1]. However, this assumption is only an approximation to reality. For example, transient data in steady-state measurement, instrument failure, human error or model non-linearity can generate non-Gaussian measurement errors [7], [8]. Outliers that are far away from the expected Gaussian distribution function can give rise to misleading estimation results. Robust estimators with non-quadratic cost functions have been introduced to solve the

outliers problem in PSSE [1], [9]. Robust estimation has also been used in conjunction with PMUs recently [6].

B. Literature Review

The existing robust dynamic state estimators can be broadly divided into two categories, one based on extensions of the Kalman filter [10] and the other based on the use of M-estimation algorithms in conjunction with the system innovation model [11], [12]. The Kalman filter based methods are intuitive. The covariance for the state estimation error for these methods, although only approximate when robust features are added to the estimators, can be obtained naturally during the estimation process. However, they cannot handle large changes in the load and power generation [13]. The M-estimation based methods are effective in the presence of outliers [14], [15]. But they can hardly be implemented in a recursive manner [16], thus generally suffer from computational load problem because of batch implementation. However, among all the M-estimators the Least Absolute Value (LAV) estimator can be made computationally efficient by transforming into a linear programming problem. Therefore, it has been commonly used in smart grids [1], [6].

The covariance of the state estimation error for the M-estimators in a dynamic system cannot be obtained easily. Previously in [16], [17] analytical equations for the covariance of the state estimation error of the M-estimator in a static system has been derived. The extension of the analytical covariance equation in [16], [17] to dynamic state estimation is not obvious. Unlike a static system, the process noise and the process matrix in a dynamic system are not assumed zero and the identity matrix respectively [4]. An example (Example 1) will be given to show that the covariance equation of a static system applied to a dynamic system gives large error. Monte-Carlo experiments can also be used to determine the covariance, but many data points are needed and hence many runs are required to achieve convergence [18]. For the IEEE 14-bus and 30-bus systems, the covariance equation derived in this paper is at least four-order of magnitude faster than using a 10,000-run Monte-Carlo experiment to obtain the covariance. It is argued that non-jacobian-based derivative free techniques based on Radial Basis Function Neural Networks (RBFNNs) [19]–[25] can also be used in robust state estimation, and can be efficient. In the case where RBFNNs methods are extended to state estimation, Monte-Carlo experiments are still required to determine the covariance. Many data points are needed and hence many runs are required to achieve convergence [18], unlike the covariance formula derived in this paper which will only require one "run" of the equation.

C. Contribution

In this paper, an analytical equation to approximately calculate the covariance of the state estimation error of the M-estimator for a dynamic system is derived. Influence Function (IF), a robust statistic tool [15], is used in the derivation of the equation. The derived analytical equation is useful. The equation gives insights into the precision of the estimation. Using the equation, the covariance of the state estimation can be expressed as functions of measurement noise variances enabling the selection of sensors for a specified estimator precision. The derived equation can be used for LAV estimator in the presence of the standard Gaussian measurement noise or Gaussian noise plus measurement outliers. Although numerical methods can also be used to find the covariance, the equation derived in this paper as a mathematical function is not only computationally efficient but also more insightful than just a numerical answer.

The paper is organized as follows. Robust dynamic state estimation problem and the LAV estimator are introduced in Section II. Influence Function analysis is given in Section III. Simulation examples and conclusions are given in Section IV and Section V respectively.

II. ROBUST DYNAMIC STATE ESTIMATION

Robust state estimation in dynamic power systems has been discussed in [6], [11]. This section only gives the equations necessary for the derivation of the covariance results in this paper. Detailed explanation of the robust state estimators as well as their stability and robustness analysis can be found in textbooks [1], [15]. The following model is widely used in the study of dynamic power systems [4], [26].

$$\begin{align} x(k+1) &= Fx(k) + Gu(k) + w(k) & (1)\\ z(k) &= Hx(k) + v(k) & (2) \end{align}$$

where $x(k) \in \mathbb{R}^{n \times 1}$, $z(k) \in \mathbb{R}^{m \times 1}$ and $u(k)$ are the state vector, PMU measurement and control signal respectively. Polar coordinates are traditionally used in Supervisory Control And Data Acquisition (SCADA) systems. In developing the problem formulation, the paper uses rectangular coordinates, where the real and imaginary parts of bus voltage phasors are used as the system states in $x(k)$ as shown in (1) and (2). Rectangular coordinates are commonly used in conjunction with PMU measurements [6], [16], [17] recently as the PMU measurements $z(k)$ can be formulated as a linear function of the state vector $x(k)$ (see equation (2)) simplifying the state estimation problem. The transformation between rectangular and polar coordinates can be found in [1], [6]. The process, input and output matrices are given by F, G and H respectively. Matrices F and G mainly depend on the setup of the system, including the system topology, the utilized devices and the operation status of the system. Matrix H mainly depends on the installation of PMUs and line impedance of the system. The formulation of (1) and (2) can be found in textbooks such as [1]. The process noise, $w(k)$, and measurement noise, $v(k)$, are stationary and independent Gaussian noise with the covariance matrices Q and R respectively. The sampling instance $k = 1, ..., N$.

Unlike a static system, the process noise and the process matrix in a dynamic system are not assumed zero and the identity matrix respectively [4].

A. The Innovation Model

The use of the innovation model in PSSE can be found in [27]. Innovation model (also referred as innovation representation) is an alternative representation of the original model and can be derived from the Kalman predictor. Two models are equivalent in the sense that their outputs have identical mean and covariance functions. The details of obtaining the innovation model as well as its properties can be found elsewhere [28, Sections 2.5, 2.6] and only the final results are given here. The model of (1) and (2) can be transformed into the innovation model.

$$x'(k+1) = Fx'(k) + Gu(k) + \Gamma\epsilon(k) \quad (3)$$
$$z(k) = Hx'(k) + \epsilon(k) \quad (4)$$

where $x'(k)$ is the state vector in the innovation model and $\epsilon(k) = z(k) - Hx'(k)$ is the innovation vector. In (3),

$$\Gamma = FK_\infty \quad (5)$$

where the steady-state Kalman gain

$$K_\infty = PH^T(HPH^T + R)^{-1} \quad (6)$$

Defining

$$\Delta x(k) = x'(k) - x(k) \quad (7)$$

the covariance is given by

$$\text{Cov}(\Delta x(k)) = P \quad (8)$$

where P can be obtained from the following Algebraic Riccati Equation

$$P = FPF^T - FPH^T(HPH^T + R)^{-1}HPF^T + Q$$

The covariance matrix of innovation vector is given as

$$\text{Cov}(\epsilon(k)) = HPH^T + R \quad (9)$$

From (9), the standard deviation of $\epsilon_i(k)$, namely $\sigma_i = \sqrt{(\text{Cov}(\epsilon(k)))_{ii}}$, and the correlation of $\epsilon_i(k)$ and $\epsilon_j(k)$, namely $\rho_{ij} = \frac{1}{\sigma_i\sigma_j}(\text{Cov}(\epsilon(k)))_{ij}$, can be obtained, where $(\cdot)_{ij}$ denotes the ith row j^{th} column element of a matrix.

B. Robust Estimator

Substituting (4) into (3) gives

$$x'(k+1) = \Phi x'(k) + Gu(k) + \Gamma z(k) \quad (10)$$

where $\Phi = F - \Gamma H$ which in practice is nonsingular.

Iterating (10) gives

$$x'(2) = \Phi x'(1) + Gu(1) + \Gamma z(1)$$
$$\vdots$$
$$x'(k) = \Phi^{k-1}x'(1) + \bar{x}(k) \quad (11)$$
$$\vdots$$
$$x'(N) = \Phi^{N-1}x'(1) + \bar{x}(N) \quad (12)$$

where

$$\bar{x}(k) = \begin{cases} 0 & k = 1 \\ \sum_{j=1}^{k-1}\Phi^{j-1}Gu(k-j) \\ + \sum_{j=1}^{k-1}\Phi^{j-1}\Gamma z(k-j) & k = 2, ..., N \end{cases} \quad (13)$$

The M-estimator can be obtained by minimizing the following cost function [1]

$$J = \sum_{i=1}^{m}\sum_{k=1}^{N}\gamma(e_i(k)) \quad (14)$$

The measurement residual

$$e_i(k) = z_i(k) - H_i\hat{x}'(k) \quad (15)$$

where $(\cdot)_i$ denotes the ith row of a matrix and $\hat{x}'(k)$ is the estimate of $x'(k)$. The cost function for the LAV estimator is well known as $\gamma = |e_i(k)|$. It is given in [1] that the LAV estimator can also be formulated as the limiting case of the Multiple-Segment estimator as follows.

$$\gamma(e_i(k)) = \begin{cases} e_i(k) - \frac{a_i^2}{2} & a_i^2 < e_i(k) \\ \frac{e_i(k)^2}{2a_i^2} & -a_i^2 \leq e_i(k) \leq a_i^2 \\ -e_i(k) - \frac{a_i^2}{2} & e_i(k) < -a_i^2 \end{cases} \quad (16)$$

where $a_i \to 0$.

C. The Least Absolute Value Estimator

Differentiating the cost function (14) of LAV estimator with respect to $\hat{x}'(N)$ gives

$$\Psi = \frac{\partial J}{\partial \hat{x}'(N)} = \sum_{i=1}^{m}\sum_{k=1}^{N}\frac{\partial\gamma(e_i(k))}{\partial(e_i(k))}\frac{\partial(e_i(k))}{\partial\hat{x}'(N)}$$
$$= -\sum_{i=1}^{m}\sum_{k=1}^{N}\frac{\partial\gamma(e_i(k))}{\partial(e_i(k))}(H_i\Phi^{k-N})^T \quad (17)$$

As $a_i \to 0$, $\frac{\partial\gamma(e_i(k))}{\partial e_i(k)}$ reduces to the sign function $\text{sgn}(e_i(k))$ and (17) becomes

$$\Psi = -\sum_{i=1}^{m}\sum_{k=1}^{N}\text{sgn}(e_i(k))(H_i\Phi^{k-N})^T \quad (18)$$

Putting (18) into matrix form using (13) and (15) gives

$$\Psi = -\tilde{H}^T\text{sgn}(E)$$
$$= -\tilde{H}^T\text{sgn}\left(Z - \tilde{H}\left(\hat{x}'(N) - \bar{x}(N)\right)\right) \quad (19)$$

where $E = [e(1)^T, ..., e(N)^T]^T$ and

$$\tilde{H} = \begin{bmatrix} H\Phi^{-N+1} \\ \vdots \\ H\Phi^0 \end{bmatrix}, \quad Z = \begin{bmatrix} z(1) - H\bar{x}(1) \\ \vdots \\ z(N) - H\bar{x}(N) \end{bmatrix} \quad (20)$$

The estimate $\hat{x}'(N)$ can be obtained by setting $\Psi = 0$. By transforming into a linear programming problem, the LAV estimator can be made computationally efficient and the details can be found in [1], [6].

III. INFLUENCE FUNCTION ANALYSIS OF LAV ESTIMATION

Using IF, an equation is derived to calculate the covariance of the state estimation error approximately.

A. Influence Function

Recently the use of IF has become popular in PSSE. For example, in [16], [17] analytical equations for the covariance of M-estimators in a static system using PMU measurements has been derived. In [10] a robust Kalman filter based estimator using PMU measurements is proposed where IF is used in an intermediate step to estimate the covariance of a priori estimation. In [14], IF of a t-distribution based maximum likelihood estimator is proposed to detect and reject bad data in PMU measurements. Further details of the IF has been derived elsewhere [10], [15], [29]. This section only gives the equations necessary for the derivation of the state estimation error covariance.

In (19), Ψ is a function of $\hat{x}'(N)$. Since LAV estimator is a unbiased estimator [15], the expectation of the estimate $\hat{x}'(N)$ is $x'(N)$. Thus, we can use the first-order Taylor serious expansion of Ψ at $\hat{x}'(N) = x'(N)$ to approximate the estimate as follows.

$$\left.\frac{\partial \Psi}{\partial \hat{x}'(N)}\right|_{\hat{x}'(N)=x'(N)} (\hat{x}'(N) - x'(N))$$
$$\approx 0 - \left.\Psi\right|_{\hat{x}'(N)=x'(N)}$$

Therefore

$$\hat{x}'(N) - x'(N)$$
$$\approx -\left[\left.\frac{\partial \Psi}{\partial \hat{x}'(N)}\right|_{\hat{x}'(N)=x'(N)}\right]^{-1} \left.\Psi\right|_{\hat{x}'(N)=x'(N)}$$

Replacing $\left.\frac{\partial \Psi}{\partial \hat{x}'(N)}\right|_{\hat{x}'(N)=x'(N)}$ by its expectation $\int_{-\infty}^{\infty} \left.\frac{\partial \Psi}{\partial \hat{x}'(N)}\right|_{\hat{x}'(N)=x'(N)} f(\epsilon)d\epsilon$ gives

$$\tilde{x}'(N) = \hat{x}'(N) - x'(N)$$
$$\approx -\left[\int_{-\infty}^{\infty} \left.\frac{\partial \Psi}{\partial \hat{x}'(N)}\right|_{\hat{x}'(N)=x'(N)} f(\epsilon)d\epsilon\right]^{-1}$$
$$\times \left.\Psi\right|_{\hat{x}'(N)=x'(N)}$$
$$= \mathrm{IF}(\epsilon) \tag{21}$$

where $\epsilon = [\epsilon(1)^T, ..., \epsilon(N)^T]^T$ and $f(\epsilon)$ is the joint probability density function (pdf) of ϵ. Equation (21) is known as the IF of the estimator [15].

In (21), $\left.\Psi\right|_{\hat{x}'(N)=x'(N)}$ and $\left.\frac{\partial \Psi}{\partial \hat{x}'(N)}\right|_{\hat{x}'(N)=x'(N)}$ are obtained as follows. From (4), (11), (12) and (19),

$$\epsilon = Z - \tilde{H}\left(x'(N) - \bar{x}(N)\right) \tag{22}$$

Using (4), (19) and (22) to obtain

$$\left.\Psi\right|_{\hat{x}'(N)=x'(N)} = -\tilde{H}^T \mathrm{sgn}(\epsilon) \tag{23}$$

From (19),

$$E = Z - \tilde{H}\left(\hat{x}'(N) - \bar{x}(N)\right) \tag{24}$$

Using (19) and (24),

$$\left.\frac{\partial \Psi}{\partial \hat{x}'(N)}\right|_{\hat{x}'(N)=x'(N)} = \left.\left(\frac{\partial \Psi}{\partial E}\frac{\partial E}{\partial \hat{x}'(N)}\right)\right|_{\hat{x}'(N)=x'(N)}$$
$$= \tilde{H}^T \frac{\partial \mathrm{sgn}(\epsilon)}{\partial \epsilon} \tilde{H} \tag{25}$$

Notice that using (22) and (24), the differential $\left.\frac{\partial \mathrm{sgn}(E)}{\partial E}\right|_{\hat{x}'(N)=x'(N)} = \frac{\partial \mathrm{sgn}(\epsilon)}{\partial \epsilon} = 2\delta$ gives the Dirac delta function point mass at $\epsilon = 0$ with the energy of 2 [30].

Substituting (23) and (25) into (21) gives

$$\tilde{x}'(N) \approx -\left[\tilde{H}^T \left(\int_{-\infty}^{\infty} \frac{\partial \mathrm{sgn}(\epsilon)}{\partial \epsilon} f(\epsilon)d\epsilon\right) \tilde{H}\right]^{-1}$$
$$\times \left(-\tilde{H}^T \mathrm{sgn}(\epsilon)\right)$$
$$= [\tilde{H}^T \Omega \tilde{H}]^{-1} \tilde{H}^T \mathrm{sgn}(\epsilon) \tag{26}$$

where

$$\Omega = \int_{-\infty}^{\infty} \frac{\partial \mathrm{sgn}(\epsilon)}{\partial \epsilon} f(\epsilon)d\epsilon$$

Notice that $\epsilon(k)$ is independent of $\epsilon(l)$, i.e. $E[\epsilon(k)\epsilon(l)] = 0$ for $k \neq l$ where $E[\cdot]$ denotes the mathematical expectation. Therefore, Ω is a diagonal matrix of dimension $mN \times mN$ given as

$$\Omega = \mathrm{diag}\left(\underbrace{\sqrt{\frac{2}{\pi\sigma_1^2}}, ..., \sqrt{\frac{2}{\pi\sigma_m^2}}, ..., \sqrt{\frac{2}{\pi\sigma_1^2}}, ..., \sqrt{\frac{2}{\pi\sigma_m^2}}}_{mN}\right) \tag{27}$$

where σ_i is given in (9).

B. Covariance Analysis

Define the state estimation error $\tilde{x}(N) = \hat{x}'(N) - x(N)$. Using (7), (8) and (21),

$$\mathrm{Cov}(\tilde{x}(N)) = \mathrm{Cov}(\tilde{x}'(N) + \Delta x(N))$$
$$= \mathrm{Cov}(\tilde{x}'(N)) + \mathrm{E}\left(\tilde{x}'(N)\Delta x(N)^T\right)$$
$$+ \mathrm{E}\left(\Delta x(N)\tilde{x}'(N)^T\right) + \mathrm{Cov}(\Delta x(N))$$
$$= \mathrm{Cov}(\tilde{x}'(N)) + \mathrm{E}\left(\tilde{x}'(N)\Delta x(N)^T\right)$$
$$+ \mathrm{E}\left(\Delta x(N)\tilde{x}'(N)^T\right) + P \tag{28}$$

The first three terms in covariance Eq.(28) can be obtained as follows.

(i) Consider the first term in covariance Eq.(28): $\mathrm{Cov}(\tilde{x}'(N))$.

Using (26),

$$
\begin{aligned}
\text{Cov} \quad (\tilde{x}'(N)) &= \int_{-\infty}^{\infty} \tilde{x}'(N)\tilde{x}'(N)^T f(\epsilon)d\epsilon \\
&\approx [\tilde{H}^T\Omega\tilde{H}]^{-1}\left[\int_{-\infty}^{\infty}\Psi\Psi^T f(\epsilon)d\epsilon\right][\tilde{H}^T\Omega\tilde{H}]^{-T} \\
&= [\tilde{H}^T\Omega\tilde{H}]^{-1} \\
&\quad \times \left[\tilde{H}^T\left(\int_{-\infty}^{\infty}\text{sgn}(\epsilon)\text{sgn}(\epsilon)^T f(\epsilon)d\epsilon\right)\tilde{H}\right][\tilde{H}^T\Omega\tilde{H}]^{-T} \\
&= [\tilde{H}^T\Omega\tilde{H}]^{-1}[\tilde{H}^T\Lambda\tilde{H}][\tilde{H}^T\Omega\tilde{H}]^{-T}
\end{aligned}
\tag{29}
$$

where $\tilde{H}$ is given in (20), Ω in (27) and

$$
\Lambda = \int_{-\infty}^{\infty}\text{sgn}(\epsilon)\text{sgn}(\epsilon)^T f(\epsilon)d\epsilon\,\text{diag}(\underbrace{\ldots \quad \Lambda^{\text{sub}} \quad \ldots}_{N})
\tag{30}
$$

$$
\Lambda^{\text{sub}} = \begin{bmatrix}
1 & \zeta_{1,2} & \cdots & \cdots & \zeta_{1,m} \\
\zeta_{2,1} & 1 & \ddots & & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\vdots & & \ddots & 1 & \zeta_{m-1,m} \\
\zeta_{m,1} & \cdots & \cdots & \zeta_{m,m-1} & 1
\end{bmatrix}
$$

$$
\begin{aligned}
\zeta_{ij} &= \int_{-\infty}^{\infty}\text{sgn}(\epsilon_i(k))\text{sgn}(\epsilon_j(k)) \\
&\quad \times f_{ij}\left(\epsilon_i(k),\epsilon_j(k)\right)d\epsilon_i(k)d\epsilon_j(k) \\
&= \frac{2}{\pi}\arcsin(\rho_{ij})
\end{aligned}
\tag{31}
$$

The function $f_{ij}\left(\epsilon_i(k),\epsilon_j(k)\right)$ is the joint pdf of $\epsilon_i(k)$ and $\epsilon_j(k)$ and ρ_{ij} can be obtained from (9). The integration results in (31) are given in [31].

(ii) Consider the second term in covariance Eq.(28): $\text{E}\left(\tilde{x}'(N)\Delta x(N)^T\right)$.

Using (2) and (4),

$$
\epsilon(k) = -H\Delta x(k) + v(k)
\tag{32}
$$

Subtracting (1) from (3) gives

$$
\Delta x(k+1) = \Phi\Delta x(k) + \Gamma v(k) - w(k)
\tag{33}
$$

From (26),

$$
\begin{aligned}
&\text{E}\left(\tilde{x}'(N)\Delta x(N)^T\right) \\
&\approx [\tilde{H}^T\Omega\tilde{H}]^{-1}\tilde{H}^T
\begin{bmatrix}
\text{E}(\text{sgn}(\epsilon(1))\Delta x(N)^T) \\
\vdots \\
\text{E}(\text{sgn}(\epsilon(k))\Delta x(N)^T) \\
\vdots \\
\text{E}(\text{sgn}(\epsilon(N))\Delta x(N)^T)
\end{bmatrix}
\end{aligned}
\tag{34}
$$

In (34), $\text{E}(\text{sgn}(\epsilon(k))\Delta x(N)^T)$ is of dimension $m\times n$, whose analytical solution is given as follows.

$$
\begin{aligned}
&\text{E}(\text{sgn}(\epsilon(k))\Delta x(N)^T) \\
&= \begin{cases} A(\Phi^{N-k})^T + B\Gamma^T(\Phi^{N-k-1})^T & k < N \\ A & k = N \end{cases}
\end{aligned}
\tag{35}
$$

where A and B are $m\times n$ and $m\times m$ matrices respectively. The elements are given as

$$
A_{ij} = \text{sgn}(-H_iP_j^T)\sqrt{\frac{2(H_iP_j^T)^2}{\pi\left(H_iPH_i^T+R_{ii}\right)}}
\tag{36}
$$

$$
B_{ij} = \begin{cases} \sqrt{\frac{2R_{ii}^2}{\pi\left(H_iPH_i^T+R_{ii}\right)}} & i = j \\ 0 & i \neq j \end{cases}
\tag{37}
$$

The derivation of (35) is given in Appendix.

(iii) Consider the third term in covariance Eq.(28): $\text{E}\left(\Delta x(N)\tilde{x}'(N)^T\right)$.

$$
\text{E}\left(\Delta x(N)\tilde{x}'(N)^T\right) = \text{E}\left(\tilde{x}'(N)\Delta x(N)^T\right)^T
\tag{38}
$$

where $\text{E}\left(\tilde{x}'(N)\Delta x(N)^T\right)$ is found in (34).

IV. Examples

In Example 1, a simple example is used to illustrate the detail calculations for covariance Eq.(28). State estimations on the IEEE 14-bus, 30-bus and 118-bus systems are given in Examples 2, 3, 4 and 5. All measurements in the examples are collected by PMUs. For easy reference, the parameters of the models and PMU measurement variances are summarized in Table I. Table I shows the precision of the PMU measurements in terms of their variance. In all the examples, the results of covariance Eq.(28) are compared with the results from Monte-Carlo experiments. It can be shown that the covariance Eq.(28) is not affected by G and $u(k)$. The effect of G and $u(k)$ is canceled out in (7) and (21). This is because the control term $Gu(k)$ is assumed to be known exactly and does not affect the precision of the state estimation [32].

A. Example 1: 2 States 4 Measurements

In (1) and (2), let

$$
F = \begin{bmatrix} 0.98 & 0 \\ 0 & 0.98 \end{bmatrix}, G = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}
\tag{39}
$$

The covariance matrices of the process noise $w(k)$ and the measurement noise $v(k)$ are given as $Q = 0.001^2 I$ and $R = \text{diag}(0.003^2, 0.004^2, 0.005^2, 0.005^2)$ respectively. In this example, the estimation batch size $N = 3$, number of measurements $m = 4$.

The innovation model is given by (3) and (4) where P in (8) is calculated as

$$
P = \begin{bmatrix} 2.916 & -0.610 \\ -0.610 & 3.287 \end{bmatrix} \times 10^{-6}
\tag{40}
$$

The covariance of innovation vector is calculated using (9),

$$
\text{Cov}(\epsilon(k)) = \begin{bmatrix} 1.192 & -0.061 & 0.231 & 0.231 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \times 10^{-5}
\tag{41}
$$

TABLE I
PARAMETERS OF THE MODELS AND NOISES USED IN THE EXAMPLES

Examples	1	2	3	4	5
Bus System	2 States 4 Measurements	IEEE 14-bus System		IEEE 30-bus System	IEEE 118-bus System
Process Model	$F = 0.98I$, $G = 0$				
Process Noise	$Q = 0.001^2 I$	$Q = 0.0001^2 I$			
Measurement Noise	$R_{11} = 0.003^2$, $R_{22} = 0.004^2$ $R_{33} = 0.005^2$ $R_{44} = 0.005^2$	$R_{ii} = 0.006^2$, $1 \le i \le 12$; $R_{ii} = 0.003^2$, $13 \le i \le 58$	Eq.(55) and Eq.(56) (with outliers)	$R_{ii} = 0.005^2$ for all measurements	
Estimator	LAV		LAV and WLS	LAV	
Figure	—	1	—	2	—
Table	—	II		III	IV

From (5), (6) and (10),

$$\Phi = F - \Gamma H = \begin{bmatrix} 0.656 & -0.067 \\ -0.067 & 0.697 \end{bmatrix} \quad (42)$$

Covariance Eq.(28) is used to calculate $\text{Cov}(\tilde{x}(3))$ and its first three terms can be obtained as follows.

(i) Consider $\text{Cov}(\tilde{x}'(3))$. Substituting (39) and (42) into (19) gives

$$\tilde{H} = \begin{bmatrix} 2.391 & 0.446 & \ldots & 1 & 1 \\ 0.446 & 2.120 & \ldots & 1 & 1 \end{bmatrix}^T \quad (43)$$

Equation (41) gives $\sigma_1^2 = 1.1916 \times 10^{-5}$, $\sigma_2^2 = 1.9286 \times 10^{-5}$, $\sigma_3^2 = \sigma_4^2 = 2.9981 \times 10^{-5}$ and upon substituting into (27) gives

$$\Omega = \text{diag}(231.1, 181.7, ..., 145.7, 145.7) \quad (44)$$

Equation (41) gives $\rho_{1,2} = -0.0402$, $\rho_{1,3} = \rho_{1,4} = 0.1220$, $\rho_{2,3} = \rho_{2,4} = 0.1113$, $\rho_{3,4} = 0.1661$ and upon substituting into (30) gives

$$\Lambda = \text{diag}(\Lambda^{\text{sub}}, \Lambda^{\text{sub}}, \Lambda^{\text{sub}}) \quad (45)$$

$$\Lambda^{\text{sub}} = \begin{bmatrix} 1 & -0.026 & 0.078 & 0.078 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

Using (29), (43), (44) and (45)

$$\text{Cov}(\tilde{x}'(3)) \approx \begin{bmatrix} 2.050 & -1.569 \\ -1.569 & 2.909 \end{bmatrix} \times 10^{-6} \quad (46)$$

(ii) Consider $\text{E}\left(\tilde{x}'(3)\Delta x(3)^T\right)$. From (34),

$$\text{E}\left(\tilde{x}'(3)\Delta x(3)^T\right)$$
$$= [\tilde{H}^T \Omega \tilde{H}]^{-1} \tilde{H}^T \begin{bmatrix} \text{E}(\text{sgn}(\epsilon(1))\Delta x(3)^T) \\ \text{E}(\text{sgn}(\epsilon(2))\Delta x(3)^T) \\ \text{E}(\text{sgn}(\epsilon(3))\Delta x(3)^T) \end{bmatrix} \quad (47)$$

In (47), $\text{E}(\text{sgn}(\epsilon(k))\Delta x(3)^T)$ is calculated using (35), where A and B are calculated using (36) and (37) respectively.

$$A = \begin{bmatrix} -0.674 & 0.111 & -0.336 & -0.336 \\ 0.141 & -0.597 & -0.390 & -0.390 \end{bmatrix}^T \times 10^{-3} \quad (48)$$

$$B = \begin{bmatrix} 2.080 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \times 10^{-3} \quad (49)$$

Substitute (48) and (49) into (35) to calculate $\text{E}(\text{sgn}(\epsilon(k))\Delta x(3)^T)$. Substituting (43), (44) and $\text{E}(\text{sgn}(\epsilon(k))\Delta x(3)^T)$ into (47) gives

$$\text{E}\left(\tilde{x}'(3)\Delta x(3)^T\right) = \begin{bmatrix} -0.353 & 0.213 \\ 0.171 & -0.494 \end{bmatrix} \times 10^{-6} \quad (50)$$

(iii) Consider $\text{E}\left(\Delta x(3)\tilde{x}'(3)^T\right)$. Using (38)

$$\text{E}\left(\Delta x(3)^T \tilde{x}'(3)^T\right) = \text{E}\left(\tilde{x}'(3)\Delta x(3)^T\right)^T$$
$$= \begin{bmatrix} -0.353 & 0.171 \\ 0.213 & -0.494 \end{bmatrix} \times 10^{-6} \quad (51)$$

Finally, substituting (40), (46), (50) and (51) into covariance Eq.(28) gives

$$\text{Cov}(\tilde{x}(3)) \approx \begin{bmatrix} 4.260 & -1.796 \\ -1.796 & 5.207 \end{bmatrix} \times 10^{-6} \quad (52)$$

The 10,000-run Monte-Carlo experiment gives

$$\text{Cov}(\tilde{x}(3)) = \begin{bmatrix} 4.397 & -1.821 \\ -1.821 & 5.162 \end{bmatrix} \times 10^{-6}$$

which is close to the result in (52).

If the covariance formula in [16], [17] is used to calculate the covariance of the state estimation error, the result is given as

$$\text{Cov}(\tilde{x}(3)) \approx \begin{bmatrix} 3.581 & -2.011 \\ -2.011 & 4.803 \end{bmatrix} \times 10^{-6}$$

which is far away from the Monte-Carlo experiment result. This is because the dynamics of the system are neglected in [16], [17], i.e. the process matrix F is assumed to be the identity matrix and the process noise $w(k)$ is not taken into account.

Notice that (52) is obtained using covariance Eq.(28) by choosing $N = 3$. For $N = 4$ and 20, covariance Eq.(28) gives

$$\text{Cov}(\tilde{x}(4)) \approx \begin{bmatrix} 3.564 & -1.247 \\ -1.247 & 4.242 \end{bmatrix} \times 10^{-6} \quad (53)$$

$$\text{Cov}(\tilde{x}(20)) \approx \begin{bmatrix} 2.916 & -0.610 \\ -0.610 & 3.286 \end{bmatrix} \times 10^{-6} \quad (54)$$

The 10,000-run Monte-Carlo experiments give

$$\text{Cov}(\tilde{x}(4)) \approx \begin{bmatrix} 3.581 & -1.219 \\ -1.219 & 4.269 \end{bmatrix} \times 10^{-6}$$

$$\text{Cov}(\tilde{x}(20)) \approx \begin{bmatrix} 2.893 & -0.634 \\ -0.634 & 3.256 \end{bmatrix} \times 10^{-6}$$

It is clear that the results in (53) and (54) are close to the Monte-Carlo experiments results.

B. Example 2: The IEEE 14-bus System

Consider the same IEEE 14-bus dynamic power system in [4] where the system process matrix $F = \text{diag}(..., 0.98, ...)$ and the process noise $w_i(k) \sim \mathcal{N}(0, 0.0001^2)$ for $i = 1, ..., 28$. The PMUs are located at buses number 2, 4, 6, 7, 9 and 13 according to [16], [33]. The standard deviation of the 12 voltage measurements and 46 current measurements are given by $v_i(k) \sim \mathcal{N}(0, 0.006^2)$ for $i = 1, ..., 12$ and $v_i(k) \sim \mathcal{N}(0, 0.003^2)$ for $i = 13, ..., m$ where $m = 58$. The state vector of the system is given by $x = [V_1^{\text{re}}, V_1^{\text{im}} ..., V_{14}^{\text{re}}, V_{14}^{\text{im}}]^T$. The real and imaginary part of the ith bus voltage phasor are given by V_i^{re} and V_i^{im} respectively. The estimation batch size $N = 3$.

The variance of the LAV estimation error calculated from covariance Eq.(28) is given in Table II Column (a). To give an idea of its accuracy, it is compared with the variance obtained from the 10,000-run Monte-Carlo experiment result in Column (b). The percentage errors between Columns (a) and (b) are given in Column (c). The maximum percentage error is 1.7%. According to the Law of Large Numbers, the Monte-Carlo experiment result becomes more accurate as more runs are involved [34]. In practice, the number of runs can be chosen based on the required accuracy of Monte-Carlo experiment result.

Covariance Eq.(28) can also be used to obtain the variance of the state estimation error as functions of measurement variances. Fig. 1 shows how the variance of the real part of Bus 1 voltage phasor, $\text{Var}(\tilde{V}_1^{\text{re}}(N))$, varies with the variance of the ith measurement, R_{ii}, $i = 1, ..., 58$. The solid curve is $\text{Var}(\tilde{V}_1^{\text{re}}(N))$ versus $R_{1,1}$ with the rest of the elements in the covariance matrix R unchanged. If for the state estimate $\text{Var}(\tilde{V}_1^{\text{re}}(N)) \leq 2.00 \times 10^{-6}$ is specified then from the solid curve, $R_{1,1} = 0.002^2$ can be selected. The crosses in the figure are variances obtained from 10,000-run Monte-Carlo experiments and they are well approximated by the solid curve. Although Monte-Carlo experiments can be used to obtain the covariance of $\hat{x}(N)$ many data points are needed and hence many simulation runs are required to achieve convergence. Ten thousand simulation runs using MATLAB 2018a on a Windows 10 computer configured with Intel® Core™, CPU $i7 - 4790$, 3.60GHz and 8GB RAM take 238 seconds to give a set of state covariance. On the same equipment, Eq.(28) takes less than 0.017 second to perform the same task. Covariance Eq.(28) is four-order of magnitude ($= 238/0.017$) faster than a 10,000-run Monte-Carlo experiment. Consider Fig. 1. There are 58
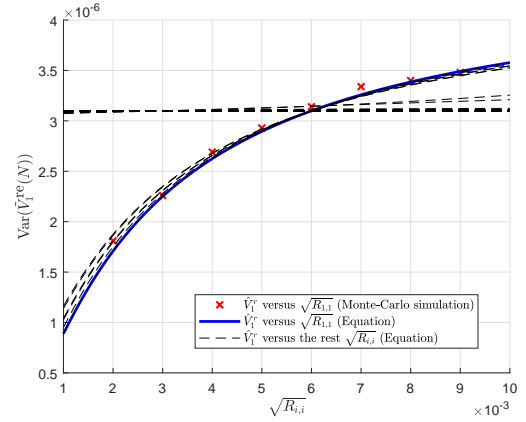

Fig. 1. $\text{Var}(\tilde{V}_1^{\text{re}}(N))$ versus $\sqrt{R_{ii}}$, $i = 1, ..., 58$. The solid curve shows $\text{Var}(\tilde{V}_1^{\text{re}}(N))$ versus $\sqrt{R_{1,1}}$. The crosses are results obtained from Monte-Carlo experiments and they matched the solid curve.

curves and each curve consists of 90 data points. A total of $90 \times 58 = 5220$ data points were used to draw Fig. 1. Eq.(28) took $0.017s \times 5220 = 1.5$ minutes to generate the data points. In contrast, Monte-Carlo experiments require $238s \times 5220 = 14.4$ days.

C. Example 3: Measurement Outliers

Instead of Gaussian measurement noise, assume that for $i = 1, ..., 12$ the measurement noise $v_i(k)$ is associated with the pdf

$$\begin{aligned} f_i(v_i(k)) = & \frac{0.99}{\sqrt{2\pi \times 0.006^2}} \exp\left(-\frac{v_i(k)^2}{2 \times 0.006^2}\right) \\ & + \frac{0.01}{\sqrt{2\pi \times 0.06^2}} \exp\left(-\frac{v_i(k)^2}{2 \times 0.06^2}\right) \end{aligned} \quad (55)$$

where the first term represents the 99% of "normal" noise while the second term represents the outliers by the 1% Gaussian noise with standard deviation (0.06) that is 10 times larger than the "normal" noise standard deviation (0.006).

For $i = 13, ..., 58$, the measurement noise $v_i(k)$ is associated with the pdf

$$\begin{aligned} f_i(v_i(k)) = & \frac{0.99}{\sqrt{2\pi \times 0.003^2}} \exp\left(-\frac{v_i(k)^2}{2 \times 0.003^2}\right) \\ & + \frac{0.01}{\sqrt{2\pi \times 0.03^2}} \exp\left(-\frac{v_i(k)^2}{2 \times 0.03^2}\right) \end{aligned} \quad (56)$$

where the first term represents the 99% of "normal" noise while the second term represents the outliers by the 1% Gaussian noise with standard deviation (0.03) that is 10 times larger than the "normal" noise standard deviation (0.003). The probability density functions in the form of a mixture distribution in (55) and (56) are also given in [16] and [15].

For uncertainties with measurement outliers, the variance of the estimation error from the Monte-Carlo experiment

TABLE II
THE VARIANCE OF THE ESTIMATION ERROR OF THE LAV AND WLS ESTIMATORS IN THE IEEE 14-BUS SYSTEM.

	(a)	(b)	(c)	(d)	(e)	(f)
Example	2			3		
Estimator	LAV					WLS
Measurement Noise	$\mathcal{N}(0, 0.006^2)$, $i = 1, ..., 12$ $\mathcal{N}(0, 0.003^2)$, $i = 13, ..., 58$			Eq.(55), $i = 1, ..., 12$ Eq.(56), $i = 13, ..., 58$		
States	Eq. (28)	Monte-Carlo	$\frac{(a)-(b)}{(b)} \times 100\%$	Monte-Carlo	$\frac{(a)-(d)}{(d)} \times 100\%$	Monte-Carlo
$\hat{V}_1^{\text{re}}$	3.10	3.14	-1.4%	3.20	-3.2%	4.00
$\hat{V}_1^{\text{im}}$	3.10	3.07	1.1%	3.20	-3.1%	4.03
$\hat{V}_2^{\text{re}}$	3.09	3.14	-1.7%	3.17	-2.5%	3.98
$\hat{V}_2^{\text{im}}$	3.09	3.06	0.9%	3.17	-2.6%	4.01
$\hat{V}_3^{\text{re}}$	3.12	3.17	-1.5%	3.18	-1.9%	4.03
$\hat{V}_3^{\text{im}}$	3.12	3.07	1.5%	3.18	-2.1%	4.02
$\hat{V}_4^{\text{re}}$	3.06	3.10	-1.5%	3.14	-2.6%	3.94
$\hat{V}_4^{\text{im}}$	3.06	3.02	1.0%	3.14	-2.7%	3.96
$\hat{V}_5^{\text{re}}$	3.06	3.11	-1.6%	3.13	-2.5%	3.94
$\hat{V}_5^{\text{im}}$	3.06	3.02	1.2%	3.14	-2.8%	3.97
$\hat{V}_6^{\text{re}}$	3.11	3.14	-0.9%	3.18	-2.1%	4.01
$\hat{V}_6^{\text{im}}$	3.11	3.09	0.7%	3.20	-2.7%	4.05
$\hat{V}_7^{\text{re}}$	3.07	3.12	-1.6%	3.16	-2.8%	3.99
$\hat{V}_7^{\text{im}}$	3.07	3.03	1.4%	3.15	-2.5%	3.95
$\hat{V}_8^{\text{re}}$	3.21	3.25	-1.1%	3.30	-2.5%	4.20
$\hat{V}_8^{\text{im}}$	3.21	3.17	1.5%	3.28	-2.1%	4.12
$\hat{V}_9^{\text{re}}$	3.08	3.12	-1.1%	3.16	-2.6%	3.99
$\hat{V}_9^{\text{im}}$	3.08	3.04	1.4%	3.17	-2.9%	3.97
$\hat{V}_{10}^{\text{re}}$	3.12	3.14	-0.6%	3.20	-2.5%	4.02
$\hat{V}_{10}^{\text{im}}$	3.12	3.08	1.2%	3.22	-3.0%	4.01
$\hat{V}_{11}^{\text{re}}$	3.34	3.36	-0.8%	3.39	-1.6%	4.34
$\hat{V}_{11}^{\text{im}}$	3.34	3.28	1.6%	3.39	-1.5%	4.35
$\hat{V}_{12}^{\text{re}}$	3.30	3.33	-1.0%	3.35	-1.4%	4.24
$\hat{V}_{12}^{\text{im}}$	3.30	3.26	1.0%	3.40	-3.0%	4.26
$\hat{V}_{13}^{\text{re}}$	3.13	3.16	-1.1%	3.20	-2.1%	4.06
$\hat{V}_{13}^{\text{im}}$	3.13	3.09	1.2%	3.22	-2.7%	4.07
$\hat{V}_{14}^{\text{re}}$	3.28	3.32	-1.5%	3.32	-1.4%	4.23
$\hat{V}_{14}^{\text{im}}$	3.28	3.23	1.6%	3.37	-2.7%	4.31
Sum of Variance (SV)	88.12	88.13	0.0%	90.32	-2.4%	114.02

Variance Unit: $\times 10^{-6}$.

is given in Table II Column (d). The measurement noise in the Monte-Carlo experiment for Column (d) is Gaussian noise plus outliers (see equations (55) and (56)). Notice that Column (a) obtained without measurement outliers in Example 2 can also be used to approximate the variance in Column (d) because the effect of the outliers are largely reduced by the robust LAV estimator. The percentage errors between Column (d) and Column (a) are given in Column (e). The maximum percentage error is 3.2%.

This example can also be used to illustrate the benefit of applying robust estimation for measurements with outliers. If the WLS estimator [1] is used instead of the LAV estimator, the variance of the state estimation error obtained from the 10,000-run Monte-Carlo experiment is given in Column (f).

D. Example 4: The IEEE 30-bus System

A new example on the IEEE 30-bus system is given as follows. The IEEE 30-bus system model is widely used in the studies of power system design and analysis [17], [35]. There are 10 PMUs strategically located at buses number 2, 4, 6, 9, 10, 12, 15, 18, 25 and 27 according to [17]. The system process matrix and process noise are chosen as $F = \text{diag}(..., 0.98, ...)$ and $w_i(k) \sim \mathcal{N}(0, 0.0001^2)$ for $i = 1, ..., 60$ as before. The the standard deviation is chosen as $v_i(k) \sim \mathcal{N}(0, 0.005^2)$ for all measurements $i = 1, ..., 104$ [17]. The estimation batch size $N = 3$. The variance of the LAV estimation error calculated from covariance Eq.(28) is given in Table III Column (a), and compared with the variance obtained from the 10,000-run Monte-Carlo experiment in Column (b). The percentage errors between Columns (a) and (b) are given in Column (c). To save space, only 2 states and the SV is presented.

The covariance Eq.(28) can be used to obtain the variation of the SV with respect to the PMU measurement variance as shown in Fig. 2. The blue curve in Fig. 2 shows that if we specify the SV of 0.9×10^{-4}, then the PMU on Bus 27 with variance of 0.002^2 can be selected with the other 9 PMUs remain unchanged. For the IEEE 30-bus system, one covariance calculation using 10,000-run Monte-Carlo experiment takes 783 seconds while equation (28) takes

TABLE III
THE VARIANCE OF THE ESTIMATION ERROR OF THE LAV ESTIMATOR IN
THE IEEE 30-BUS SYSTEM.

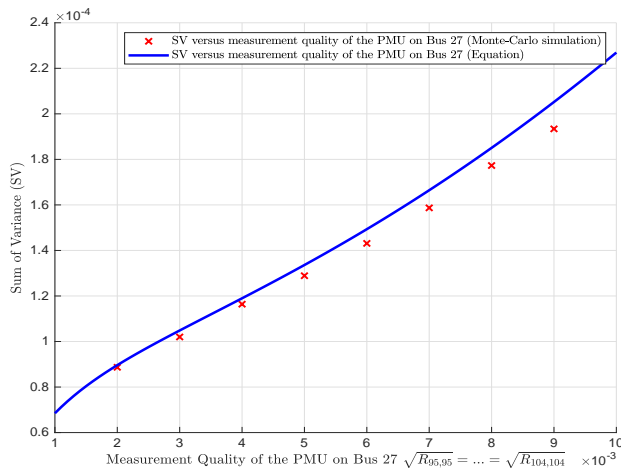|	(a)	(b)	(c)
States	Eq. (28)	Monte-Carlo	$\frac{(a)-(b)}{(b)} \times 100\%$
$\hat{V}_1^{\text{re}}$	1.50	1.52	-1.0%
$\hat{V}_1^{\text{im}}$	1.50	1.48	1.6%
$\vdots$	$\vdots$	$\vdots$	$\vdots$
SV	133.77	128.88	3.8%

Variance Unit: $\times 10^{-6}$.

Fig. 2. SV versus measurement quality of the PMU on Bus 27 while the other 9 PMUs remain unchanged. Measurement variance $R_{i,i}, i = 95, 96, ..., 104$ are associated with the PMU on Bus 27. The solid curve shows SV versus the quality of the PMU on Bus 27 calculated from covariance Eq.(28). The crosses are results obtained from Monte-Carlo experiments and they matched the solid curve.

0.065 second, four-order of magnitude faster. In example 2, using 10,000-run Monte-Carlo experiments 14.4 days are required to draw Fig. 1 for the IEEE 14-bus system. To draw a similar figure for the IEEE 30-bus system, 79.9 days would be required. Even if a lower level of precision is accepted and the Monte-Carlo runs are reduced to 1,000, it will still take 8 days.

E. *Example 5: The IEEE 118-bus System*

A larger IEEE 118-bus system is also used to verify the precision of the proposed method. The PMUs are located according to [33]. The process matrix, process noise, measurement noise and estimation batch size are chosen as $F = \text{diag}(..., 0.98, ...)$, $w_i(k) \sim \mathcal{N}(0, 0.0001^2)$ for $i = 1, ..., 236$, $v_i(k) \sim \mathcal{N}(0, 0.005^2)$ for all measurements $i = 1, ..., 452$ and $N = 3$ respectively. The variance of the LAV estimation error calculated from covariance Eq.(28) is given in Table IV Column (a), and compared with the variance obtained from the 10,000-run Monte-Carlo experiment in Column (b). The percentage errors between Columns (a) and (b) are given in Column (c). To save space, only 2 states and the SV is presented. The percentage error of SV between the proposed formula and the 10,000-

run Monte-Carlo experiment is 3.5%, which should be practically acceptable.

TABLE IV
THE VARIANCE OF THE ESTIMATION ERROR OF THE LAV ESTIMATOR IN
THE IEEE 118-BUS SYSTEM.

	(a)	(b)	(c)
States	Eq. (28)	Monte-Carlo	$\frac{(a)-(b)}{(b)} \times 100\%$
$\hat{V}_1^{\text{re}}$	0.765	0.738	3.6%
$\hat{V}_1^{\text{im}}$	0.765	0.725	5.5%
$\vdots$	$\vdots$	$\vdots$	$\vdots$
SV	163.86	158.31	3.5%

Variance Unit: $\times 10^{-6}$.

V. CONCLUSION

In this paper, an analytical equation is derived using IF approximation to calculate the covariance of the dynamic state estimation error of the robust LAV estimator. The equation gives insights into the precision of the estimation and can be used to express the variances of the state estimates as functions of measurement noise variances, enabling the selection of sensors for specified estimator precision. Simulations on the IEEE 14-bus, 30-bus and 118-bus systems are given to illustrate the usefulness of the equation. Monte-Carlo experiments can also be used to determine the covariance, but many data points are needed and hence many runs are required to achieve convergence. For the IEEE 14-bus and 30-bus examples, the covariance equation is at least four-order of magnitude faster than the 10,000-run Monte-Carlo experiment. The derived equation can be used for LAV estimator in the presence of the standard Gaussian measurement noise or Gaussian noise plus measurement outliers because the effect of the outliers are largely reduced by the robust LAV estimator. The maximum percentage error between covariance equation and the Monte-Carlo experiment result in the presence of measurement outliers is 3.2% in the IEEE 14-bus system example.

REFERENCES

[1] A. Abur and A. Gómez-Expósito, *Power system state estimation: theory and implementation*. CRC press, 2004.
[2] V. Basetti, A. K. Chandel, and R. Chandel, "Power system dynamic state estimation using prediction based evolutionary technique," *Energy*, vol. 107, pp. 29–47, 2016.
[3] A. K. Singh and B. C. Pal, "Decentralized dynamic state estimation in power systems using unscented transformation," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 794–804, 2014.
[4] L. Hu, Z. Wang, I. Rahman, and X. Liu, "A constrained optimization approach to dynamic state estimation for power systems including PMU and missing measurements," *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 2, pp. 703–710, 2016.
[5] Y.-F. Huang, S. Werner, J. Huang, N. Kashyap, and V. Gupta, "State estimation in electric power grids: Meeting new challenges presented by the requirements of the future grid," *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 33–43, 2012.
[6] M. Göl and A. Abur, "LAV based robust state estimation for systems measured by PMUs," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 1808–1814, 2014.

[7] W. K. Ho, K. V. Ling, H. D. Vu, and X. Wang, "Filtering of the ARMAX process with generalized t-distribution noise: The influence function approach," *Ind. & Eng. Chemistry Research*, vol. 53, no. 17, pp. 7019–7028, 2014.

[8] D. Wang and J. Romagnoli, "A framework for robust data reconciliation based on a generalized objective function," *Ind. & Eng. Chemistry Research*, vol. 42, no. 13, pp. 3075–3084, 2003.

[9] A. Monticelli, *State estimation in electric power systems: a generalized approach.* Springer Science & Business Media, 1999, vol. 507.

[10] J. Zhao, M. Netto, and L. Mili, "A robust iterated extended kalman filter for power system dynamic state estimation," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3205–3216, July 2017.

[11] F. Aminifar, M. Shahidehpour, M. Fotuhi-Firuzabad, and S. Kamalinia, "Power system dynamic state estimation with synchronized phasor measurements," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 2, pp. 352–363, 2014.

[12] J. Zhao and L. Mili, "A robust generalized-maximum likelihood unscented Kalman filter for power system dynamic state estimation," *IEEE J. Sel. Topics Signal Proc.*, vol. 12, no. 4, pp. 578–592, Aug 2018.

[13] J. Zhao, M. Netto, and L. Mili, "A robust iterated extended kalman filter for power system dynamic state estimation," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3205–3216, 2017.

[14] T. Chen, L. Sun, K. Ling, and W. Ho, "Robust power system state estimation using t-distribution noise model," *IEEE Syst. J.*, July doi:10.1109/JSYST.2018.2890106, in press, 2019.

[15] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust statistics: the approach based on influence functions.* John Wiley & Sons, 2011, vol. 114.

[16] W. K. Ho, T. Chen, K. V. Ling, and L. Sun, "Variance analysis of robust state estimation in power system using influence function," *Int. J. of Elect. Power & Energy Systems*, vol. 92, pp. 53–62, 2017.

[17] L. Sun, T. Chen, X. Chen, W. K. Ho, K.-V. Ling, K.-J. Tseng, and G. A. Amaratunga, "Optimum placement of phasor measurement units in power systems," *IEEE Trans. on Instrum. and Meas.*, no. 68, pp. 421–429, 2019.

[18] R. Singh, B. C. Pal, and R. B. Vinter, "Measurement placement in distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 668–675, 2009.

[19] H. R. Baghaee, M. Mirsalim, G. B. Gharehpetian, and H. A. Talebi, "Generalized three phase robust load-flow for radial and meshed power systems with and without uncertainty in energy resources using dynamic radial basis functions neural networks," *J. of cleaner production*, vol. 174, pp. 96–113, 2018.

[20] ——, "Three-phase ac/dc power-flow for balanced/unbalanced microgrids including wind/solar, droop-controlled and electronically-coupled distributed energy resources using radial basis function neural networks," *IET Power Electron.*, vol. 10, no. 3, pp. 313–328, 2017.

[21] ——, "Unbalanced harmonic power sharing and voltage compensation of microgrids using radial basis function neural network-based harmonic power-flow calculations for distributed and decentralised control structures," *IET Gener. Transm. Distrib.*, vol. 12, no. 7, pp. 1518–1530, 2017.

[22] H. R. Baghaee, M. Mirsalim, and G. B. Gharehpetian, "Power calculation using rbf neural networks to improve power sharing of hierarchical control scheme in multi-der microgrids," *IEEE Trans. Emerg. Sel. Topics Circuits Syst.*, vol. 4, no. 4, pp. 1217–1225, 2016.

[23] H. R. Baghaee, M. Mirsalim, G. B. Gharehpetian, and H. A. Talebi, "Nonlinear load sharing and voltage compensation of microgrids based on harmonic power-flow calculations using radial basis function neural networks," *IEEE Syst. J.*, vol. 12, no. 3, pp. 2749–2759, 2018.

[24] H. R. Baghaee, M. Mirsalim, G. B. Gharehpetian, and H. A. Talebi, "Fuzzy unscented transform for uncertainty quantification of correlated wind/pv microgrids: possibilistic–probabilistic power flow based on rbfnns," *IET Renew. Power Gener.*, vol. 11, no. 6, pp. 867–877, 2017.

[25] ——, "Application of rbf neural networks and unscented transformation in probabilistic power-flow of microgrids including correlated wind/pv units and plug-in hybrid electric vehicles," *Simulation Modelling Practice and Theory*, vol. 72, pp. 51–68, 2017.

[26] Z. Guo, S. Li, X. Wang, and W. Heng, "Distributed point-based gaussian approximation filtering for forecasting-aided state estimation in power systems," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2597–2608, 2016.

[27] H. M. Beides and G. Heydt, "Dynamic state estimation of power system harmonics using kalman filter methodology," *IEEE Trans. Power Del.*, vol. 6, no. 4, pp. 1663–1670, 1991.

[28] F. L. Lewis, L. Xie, and D. Popa, *Optimal and robust estimation: with an introduction to stochastic control theory.* CRC press, 2007, vol. 29.

[29] P. J. Huber, *Robust statistics.* Springer, 2011.

[30] G. B. Arfken and H. J. Weber, "Mathematical methods for physicists," 1999.

[31] Y. L. Tong, *The multivariate normal distribution.* Springer Science & Business Media, 2012.

[32] K. V. Ling and K. W. Lim, "Receding horizon recursive state estimation," *IEEE Trans. Autom. Control*, vol. 44, no. 9, pp. 1750–1753, 1999.

[33] P. Yang, Z. Tan, A. Wiesel, and A. Nehorai, "Power system state estimation using PMUs with imperfect synchronization," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4162–4172, 2013.

[34] Y. A. Shreider, *The Monte Carlo method: the method of statistical trials.* Elsevier, 2014, vol. 87.

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

APPENDIX

The derivation of (35) is given below.

$$\mathrm{E}(\mathrm{sgn}(\epsilon(k))\Delta x(N)^T) = \begin{bmatrix} \mathrm{E}(\mathrm{sgn}(\epsilon_1(k))\Delta x(N)^T) \\ \vdots \\ \mathrm{E}(\mathrm{sgn}(\epsilon_m(k))\Delta x(N)^T) \end{bmatrix} \tag{57}$$

Case 1: $k < N$.

Iteratively writing (33) gives

$$
\begin{aligned}
\Delta x(k+1) &= \Phi\Delta x(k) + \Gamma v(k) - w(k) \\
\Delta x(k+2) &= \Phi\Delta x(k+1) + \Gamma v(k+1) - w(k+1) \\
&= \Phi^2\Delta x(k) + \Phi\Gamma v(k) + \Gamma v(k+1) \\
&\quad -\Phi w(k) - w(k+1) \\
\vdots &= \vdots \\
\Delta x(N) &= \Phi^{N-k}\Delta x(k) + \Phi^{N-k-1}\Gamma v(k) \\
&\quad + \sum_{i=1}^{N-k-1} \Phi^{i-1}\Gamma v(N-i) \\
&\quad - \sum_{i=1}^{N-k} \Phi^{i-1}w(N-i)
\end{aligned}
\tag{58}
$$

Use (32) and (58) to derive (59). Notice that the last two terms $\sum_{i=1}^{N-k-1} \Phi^{i-1}\Gamma v(N-i)$ and $\sum_{i=1}^{N-k} \Phi^{i-1}w(N-i)$ in (58) are independent of $\epsilon(k)$. Therefore,

$$
\begin{aligned}
&\mathrm{E}(\mathrm{sgn}(\epsilon_i(k))\Delta x(N)^T) \\
&= \mathrm{E}\big(\mathrm{sgn}(-H_i\Delta x(k) + v_i(k)) \\
&\quad \times \big(\Delta x(k)^T(\Phi^{N-k})^T + v(k)^T\Gamma^T(\Phi^{N-k-1})^T\big)\big) \\
&= \mathrm{E}\big(\mathrm{sgn}(-H_i\Delta x(k) + v_i(k))\Delta x(k)^T\big)(\Phi^{N-k})^T \\
&\quad + \mathrm{E}\big(\mathrm{sgn}(-H_i\Delta x(k) + v_i(k))v(k)^T\big)\Gamma^T(\Phi^{N-k-1})^T \\
&= A_i(\Phi^{N-k})^T + B_i\Gamma^T(\Phi^{N-k-1})^T
\end{aligned}
\tag{59}
$$

where A_i and B_i are $1 \times n$ and $1 \times m$ row vectors respectively given by

$$
\begin{aligned}
A_i &= \mathrm{E}\left(\mathrm{sgn}(-H_i\Delta x(k) + v_i(k))\Delta x(k)^T\right) \\
&= \begin{bmatrix} \mathrm{E}\left(\mathrm{sgn}\left(-H_i\Delta x(k) + v_i(k)\right)\Delta x_1(k)\right) \\ \vdots \\ \mathrm{E}\left(\mathrm{sgn}\left(-H_i\Delta x(k) + v_i(k)\right)\Delta x_n(k)\right) \end{bmatrix}^T \quad (60)
\end{aligned}
$$

$$
\begin{aligned}
B_i &= \mathrm{E}\left(\mathrm{sgn}(-H_i\Delta x(k) + v_i(k))v(k)^T\right) \\
&= \begin{bmatrix} \mathrm{E}\left(\mathrm{sgn}\left(-H_i\Delta x(k) + v_i(k)\right)v_1(k)\right) \\ \vdots \\ \mathrm{E}\left(\mathrm{sgn}\left(-H_i\Delta x(k) + v_i(k)\right)v_m(k)\right) \end{bmatrix}^T \quad (61)
\end{aligned}
$$

In (60) and (61), A_i and B_i are calculated as follows.

Consider A_i. The jth element in A_i in (60) is given by

$$
A_{ij} = \mathrm{E}\left(\mathrm{sgn}\left(-H_i\Delta x(k) + v_i(k)\right)\Delta x_j(k)\right) \quad (62)
$$

where $-H_i\Delta x(k)$ and $\Delta x_j(k)$ are correlated. The correlation can be determined as follows.

$$
\begin{aligned}
&\mathrm{Corr}\left(-H_i\Delta x(k), \Delta x_j(k)\right) \\
&= \frac{\mathrm{Cov}\left(-H_i\Delta x(k), \Delta x_j(k)\right)}{\sqrt{\mathrm{Var}(-H_i\Delta x(k))}\sqrt{\mathrm{Var}(\Delta x_j(k))}} \quad (63)
\end{aligned}
$$

where $\mathrm{Corr}(\cdot)$ denotes correlation. In (63),

$$
\begin{aligned}
\mathrm{Cov}\left(-H_i\Delta x(k), \Delta x_j(k)\right) &= \mathrm{E}\left((-H_i\Delta x(k))\Delta x_j(k)\right) \\
&= \mathrm{E}\left(-\sum_{l=1}^{n} H_{il}\Delta x_l(k)\Delta x_j(k)\right) = -H_i P_j^T \quad (64)
\end{aligned}
$$

$$
\mathrm{Var}(-H_i\Delta x(k)) = H_i P H_i^T \quad (65)
$$

$$
\mathrm{Var}(\Delta x_j(k)) = P_{jj} \quad (66)
$$

where P is given in (8). Substituting (64), (65) and (66) into (63) gives

$$
\mathrm{Corr}\left(-H_i\Delta x(k), \Delta x_j(k)\right) = -\frac{H_i P_j^T}{\sqrt{H_i P H_i^T}\sqrt{P_{jj}}} \quad (67)
$$

Using the correlation in (67), isolate the part of $-H_i\Delta x(k)$ in (62) that is independent of $\Delta x_j(k)$ as follows.

$$
-H_i\Delta x(k) = -H_i\Delta x(k) + \frac{H_i P_j^T}{P_{jj}}\Delta x_j(k) - \frac{H_i P_j^T}{P_{jj}}\Delta x_j(k) \quad (68)
$$

where the expression in the square bracket is independent of $\Delta x_j(k)$. This can be verified by showing that the correlation between the square bracket and $\Delta x_j(k)$ is zero.

$$
\begin{aligned}
&\mathrm{Corr}\left(\left[-H_i\Delta x(k) + \frac{H_i P_j^T}{P_{jj}}\Delta x_j(k)\right], \Delta x_j(k)\right) \\
&= \frac{\mathrm{Cov}\left(\left[-H_i\Delta x(k) + \frac{H_i P_j^T}{P_{jj}}\Delta x_j(k)\right], \Delta x_j(k)\right)}{\sqrt{\mathrm{Var}\left[-H_i\Delta x(k) + \frac{H_i P_j^T}{P_{jj}}\Delta x_j(k)\right]}\mathrm{Var}(\Delta x_j(k))} \quad (69)
\end{aligned}
$$

Substituting (64), (65), and (8) into (69) gives

$$
\mathrm{Corr}\left(\left[-H_i\Delta x(k) + \frac{H_i P_j^T}{P_{jj}}\Delta x_j(k)\right], \Delta x_j(k)\right) = 0
$$

Finally, substituting (68) into (62) gives

$$
\begin{aligned}
A_{ij} &= \mathrm{E}\left(\mathrm{sgn}\left(-\frac{H_i P_j^T}{P_{jj}}\Delta x_j(k) - H_i\Delta x(k)\right.\right. \\
&\quad \left.\left. + \frac{H_i P_j^T}{P_{jj}}\Delta x_j(k) + v_i(k)\right)\Delta x_j(k)\right) \quad (70) \\
&= \mathrm{E}\left(\mathrm{sgn}\left(c_{ij}\Delta x_j(k) + \bar{v}_{ij}(k)\right)\Delta x_j(k)\right) \quad (71)
\end{aligned}
$$

where

$$
c_{ij} = -\frac{H_i P_j^T}{P_{jj}} \quad (72)
$$

$$
\bar{v}_{ij}(k) = -H_i\Delta x(k) + \frac{H_i P_j^T}{P_{jj}}\Delta x_j(k) + v_i(k)
$$

Obviously $\bar{v}_{ij}(k)$ is independent of $\Delta x_j(k)$. Denote the variance of $\bar{v}_{ij}(k)$ as

$$
\begin{aligned}
s_{\bar{v}_{ij}}^2 &= \mathrm{Var}(\bar{v}_{ij}(k)) \\
&= \mathrm{Var}(-H_i\Delta x(k)) + 2\mathrm{Cov}\left(-H_i\Delta x(k), \frac{H_i P_j^T}{P_{jj}}\Delta x_j(k)\right) \\
&\quad + \mathrm{Var}\left(\frac{H_i P_j^T}{P_{jj}}\Delta x_j(k)\right) + \mathrm{Var}(v_i(k)) \quad (73)
\end{aligned}
$$

Substituting (64), (65), (8) and $\mathrm{Var}(v_i(k)) = R_{ii}$ into (73) gives

$$
\begin{aligned}
s_{\bar{v}_{ij}}^2 &= H_i P H_i^T - 2\frac{(H_i P_j^T)^2}{P_{jj}} + \frac{(H_i P_j^T)^2}{P_{jj}} + R_{ii} \\
&= H_i P H_i^T - \frac{(H_i P_j^T)^2}{P_{jj}} + R_{ii} \quad (74)
\end{aligned}
$$

Since $\Delta x_j(k)$ and $\bar{v}_{ij}(k)$ are independent, (71) can be solved using

$$
\begin{aligned}
A_{ij} &= \iint_{-\infty}^{\infty} \mathrm{sgn}\left(c_{ij}\Delta x_j(k) + \bar{v}_{ij}(k)\right)\Delta x_j(k) \\
&\quad \times f_{\Delta x_j(k)} f_{\bar{v}_{ij}(k)} d(\Delta x_j(k)) d(\bar{v}_{ij}(k)) \quad (75)
\end{aligned}
$$

where $f_{\Delta x_j(k)}$ and $f_{\bar{v}_{ij}(k)}$ are pdfs given by

$$
f_{\Delta x_j(k)} = \frac{1}{\sqrt{2\pi P_{jj}}}\exp\left(-\frac{\Delta x_j(k)^2}{2P_{jj}}\right) \quad (76)
$$

$$
f_{\bar{v}_{ij}(k)} = \frac{1}{\sqrt{2\pi s_{\bar{v}_{ij}}^2}}\exp\left(-\frac{\bar{v}_{ij}(k)^2}{2s_{\bar{v}_{ij}}^2}\right) \quad (77)
$$

Rewriting (75) gives

$$
\begin{aligned}
A_{ij} &= \mathrm{sgn}(c_{ij})\iint_{-\infty}^{\infty}\mathrm{sgn}\left(\Delta x_j(k) + \frac{\bar{v}_{ij}(k)}{c_{ij}}\right)\Delta x_j(k) \\
&\quad \times f_{\Delta x_j(k)} f_{\bar{v}_{ij}(k)} d(\Delta x_j(k)) d(\bar{v}_{ij}(k)) \quad (78)
\end{aligned}
$$

Substitute (76) and (77) into (78). Notice that

$$\text{sgn}\left(\Delta x_j(k) + \frac{\bar{v}_{ij}(k)}{c_{ij}}\right)$$
$$= \begin{cases} -1 & -\infty \leq & \Delta x_j(k) & \leq -\bar{v}_{ij}(k)/c_{ij} \\ 1 & -\bar{v}_{ij}(k)/c_{ij} \leq & \Delta x_j(k) & \leq \infty \end{cases}$$

Therefore,

$$A_{ij} = \text{sgn}(c_{ij}) \int_{-\infty}^{\infty} f_{\bar{v}_{ij}(k)}$$
$$\times \left[\int_{-\infty}^{-\bar{v}_{ij}(k)/c_{ij}} (-1)\Delta x_j(k) f_{\Delta x_j(k)} d(\Delta x_j(k))\right.$$
$$\left. + \int_{-\bar{v}_{ij}(k)/c_{ij}}^{\infty} (+1)\Delta x_j(k) f_{\Delta x_j(k)} d(\Delta x_j(k))\right] d(\bar{v}_{ij}(k))$$
$$= \text{sgn}(c_{ij}) \int_{-\infty}^{\infty} f_{\bar{v}_{ij}(k)}$$
$$\times \left[2P_{jj}\frac{1}{\sqrt{2\pi P_{jj}}}\exp\left(-\frac{(\bar{v}_{ij}(k)/c_{ij})^2}{2P_{jj}}\right)\right] d(\bar{v}_{ij}(k))$$
$$= \text{sgn}(c_{ij}) \sqrt{\frac{2P_{jj}^2}{\pi(P_{jj} + \frac{s_{\bar{v}_{ij}}^2}{c_{ij}^2})}} \tag{79}$$

Substitute c_{ij} and $s_{\bar{v}_{ij}}^2$ from (72) and (74) into (79). Noticing that P_{jj} is always positive,

$$A_{ij} = \text{sgn}(-H_i P_j^T)\sqrt{\frac{2(H_i P_j^T)^2}{\pi\left(H_i P H_i^T + R_{ii}\right)}} \tag{80}$$

Consider B_i. The jth element in B_i in (61) is given by

$$B_{ij} = \text{E}\left(\text{sgn}\left(-H_i \Delta x(k) + v_i(k)\right) v_j(k)\right) \tag{81}$$

Let $\bar{z}_i(k) = -H_i \Delta x(k)$ whose variance is denoted by $s_{\bar{z}_i}^2$. Notice that $\bar{z}_i(k)$ is independent from $v_j(k)$ and

$$s_{\bar{z}_i}^2 = \text{Var}(-H_i \Delta x(k)) = H_i P H_i^T \tag{82}$$

Replacing $-H_i \Delta x(k)$ in (81) with $\bar{z}_i(k)$ gives

$$B_{ij} = \text{E}\left(\text{sgn}\left(\bar{z}_i(k) + v_i(k)\right) v_j(k)\right) \tag{83}$$

If $i = j$, (83) can be solved as follows. Since $v_i(k)$ and $\bar{z}_i(k)$ are independent,

$$B_{ii} = \iint_{-\infty}^{\infty} \text{sgn}\left(\bar{z}_i(k) + v_i(k)\right) v_i(k) f_{v_i(k)}$$
$$\times f_{\bar{z}_i(k)} d(v_i(k)) d(\bar{z}_i(k)) \tag{84}$$

where

$$f_{v_i(k)} = \frac{1}{\sqrt{2\pi R_{ii}}}\exp\left(-\frac{v_i(k)^2}{2R_{ii}}\right) \tag{85}$$

$$f_{\bar{z}_i(k)} = \frac{1}{\sqrt{2\pi s_{\bar{z}_i}^2}}\exp\left(-\frac{\bar{z}_i(k)^2}{2s_{\bar{z}_i}^2}\right) \tag{86}$$

Substitute (85) and (86) into (84). Notice that

$$\text{sgn}\left(v_i(k) + \bar{z}_i(k)\right)$$
$$= \begin{cases} -1 & -\infty \leq & v_i(k) & \leq -\bar{z}_i(k) \\ 1 & -\bar{z}_i(k) \leq & v_i(k) & \leq \infty \end{cases}$$

Therefore,

$$B_{ii} = \int_{-\infty}^{\infty} f_{\bar{z}_i(k)} \left[\int_{-\infty}^{-\bar{z}_i(k)} (-1)v_i(k) f_{v_i(k)} d(v_i(k))\right.$$
$$\left. + \int_{-\bar{z}_i(k)}^{\infty} (+1)v_i(k) f_{v_i(k)} d(v_i(k))\right] d(\bar{z}_i(k))$$
$$= \int_{-\infty}^{\infty} f_{\bar{z}_i(k)} \left[2R_{ii}\frac{1}{\sqrt{2\pi R_{ii}}}\exp\left(-\frac{\bar{z}_i(k)^2}{2R_{ii}}\right)\right] d(\bar{z}_i(k))$$
$$= \sqrt{\frac{2R_{ii}^2}{\pi(R_{ii} + s_{\bar{v}_{ii}}^2)}} \tag{87}$$

Substituting $s_{\bar{z}_i}^2$ in (82) into (87) gives

$$B_{ii} = \sqrt{\frac{2R_{ii}^2}{\pi\left(H_i P H_i^T + R_{ii}\right)}}$$

Finally,

$$B_{ij} = \begin{cases} \sqrt{\frac{2R_{ii}^2}{\pi\left(H_i P H_i^T + R_{ii}\right)}} & i = j \\ 0 & i \neq j \end{cases} \tag{88}$$

where from (83) $B_{ij} = 0$ for $i \neq j$ since $\bar{z}_i(k) + v_i(k)$ is independent of $v_j(k)$.

Case 2: $k = N$.

From (32), $\epsilon(N) = -H\Delta x(N) + v(N)$ Therefore, using (58) and (59)

$$\text{E}(\text{sgn}(\epsilon_i(N))\Delta x(N)^T)$$
$$= \text{E}\left(\text{sgn}(-H_i \Delta x(N) + v_i(N))\Delta x(N)^T\right) = A_i \tag{89}$$

where A_{ij} is calculated using (80).

Let $A = [..., A_i^T, ...]^T$, $B = [..., B_i^T, ...]^T$. Substituting (59) and (89) into (57) gives

$$\text{E}(\text{sgn}(\epsilon(k))\Delta x(N)^T)$$
$$= \begin{cases} A(\Phi^{N-k})^T + B\Gamma^T(\Phi^{N-k-1})^T & k < N \\ A & k = N \end{cases} \tag{90}$$

Equations (90), (80) and (88) correspond to (35), (36) and (37) respectively.

Point-by-point Response to Editor Comment.

Editor: The paper requires minor revisions as suggested by reviewer.

Response: The paper has been revised accordingly. The reply to the reviewer is given below.

Point-by-point Response to Reviewer 2 Comment.

Reviewer: Although the authors have revised this manuscript according to reviewers' comments, the size of the system (IEEE 30-bus system) is still too small for state estimations. Please conduct simulation studies of practical large-scale system such as IEEE 118-bus, IEEE 300 Bus, and Polish system as mentioned in Matpower.

Response: A new "Example 5: The IEEE 118-bus System" is added on page 9 of the revised paper as follows.

"A larger IEEE 118-bus system is also used to verify the precision of the proposed method. The PMUs are located according to [33]. The process matrix, process noise, measurement noise and estimation batch size are chosen as $F = \text{diag}(..., 0.98, ...)$, $w_i(k) \sim \mathcal{N}(0, 0.0001^2)$ for $i = 1, ..., 236$, $v_i(k) \sim \mathcal{N}(0, 0.005^2)$ for all measurements $i = 1, ..., 452$ and $N = 3$ respectively. The variance of the LAV estimation error calculated from covariance Eq.(28) is given in Table IV Column (a), and compared with the variance obtained from the 10,000-run Monte-Carlo experiment in Column (b). The percentage errors between Columns (a) and (b) are given in Column (c). To save space, only 2 states and the SV is presented. The percentage error of SV between the proposed formula and the 10,000-run Monte-Carlo experiment is 3.5%, which should be practically acceptable. "

Table IV is given on page 9 of the revised paper.