Distributed State Estimation in Smart Grid with Communication Constraints

Hang Ma, Yu-Han Yang, Yan Chen, and K. J. Ray Liu
Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742
E-mail: {hangma, yhyang, yan, kjrliu}@umd.edu

***Abstract*—Distributed state estimation in smart grid highly relies on the availability of measurements. Transmitting a lot of measurements within a small time interval is costly and sometimes even impossible. This paper explores the problem of distributed state estimation in smart grid with constraint on the number of measurements that is able to be transmitted in one step. It is shown that there exists a lower bound which depends on the structure of the grid such that if the number of permissible measurements is beyond the bound, then the estimator achieves the same performance as its peer without the constraint. Further, if the number of permissible measurements is below the lower bound, a tradeoff between the performance of the estimator and the measurements transmitted is needed to meet the constraint. A method to attain the tradeoff is offered in this paper. The proposed conclusions and methods are illustrated in the simulation on the IEEE 14-bus system.**

I. Introduction

The power grid in the United States has evolved over the past century from a series of small independent community-based systems to a large-scale and complex system involving many kinds of components. Such a system entails advanced operating methods that are more sensitive, reliable and economic than before. Efficient operation of the system requires precise real-time estimation of the states [1]. There has been a lot of research on the state estimation for large-scale systems [2]–[4].

Compared with traditional state estimation of the large-scale system, state estimation on smart grid has a lot of differences. Fortunately, we usually do not have power constraint in computation because they have access to the power. Moreover, the sensors are generally not mobile thus it is possible to use stationary grouping. However, there are some other features that make the state estimation in power grid a special problem that needs to be investigated.

One of the features is that the power grid can be distributed in quite a variety of environments, for instance some island wind farms [5] and some grid across mountains. Such environment may limit the choice of communication methods. Some of the traditional communication methods cannot be used in such environment. Moreover, transmitting high-rate measurements in such environment may be expensive by the choice of communication methods and sometimes even impossible. Thus, distributed estimation of the state is always needed to reduce the communication cost.

Another feature is that there is no existing model for the state evolution. The state of the power grid is affected by many factors such as the time of the day and the weather. There is no existing model that could take all these factors into account [6]. Thus, without the knowledge of evolution model, the choice of estimation method is also limited. There has been a lot of research on the distributed state estimation on power grid [7]–[11]. One common approach is to use the hierarchical method [7], [8], where the local measurements are first combined by the local estimator to estimate the states and then transmitted to a central coordinator for further estimation. Similar approaches are proposed by using a fully distributed way without the central coordinator [9]–[11].

In this paper, we investigate the problem of distributed state estimation in smart grid with communication constraints. It is motivated by the fact that in some cases, transmitting a lot of data in a time interval is expensive and sometimes even impossible, thus reducing the communication burden is necessary. In this paper, we define the communication capability as the number of measurements to be transmitted from the pre-processing station to the global estimation center in one time slot, which is an important criterion since it is proportional to the required bandwidth of transmission. We propose a distributed state estimation approach to reduce the number of measurements needed to be transmitted. Each substation first transmits its measurements to a local pre-processing station that will transmit the processed measurements (not necessarily the local estimation) to a global estimation center in distance. We first derive a lower bound for the communication capability such that if the communication capability is beyond the bound, then we can achieve the same performance as the global minimum mean square error (MMSE) estimator in a distributed way. When the communication capability is below the bound, we show that there exists a tradeoff between the performance of the estimator and the measurements transmitted. We further propose a method to achieve this tradeoff.

The rest of this paper is organized as follows. In section II, we describe the problem formulation. The lower bound for the communication capability is derived in section III. In section IV, we introduce in details the proposed optimal state estimator. Finally, we show simulation results in section V and draw conclusions in section VI.

II. Problem Formulation

We consider a state estimation problem on a smart grid with its substations mainly distributed in k areas as shown in Figure

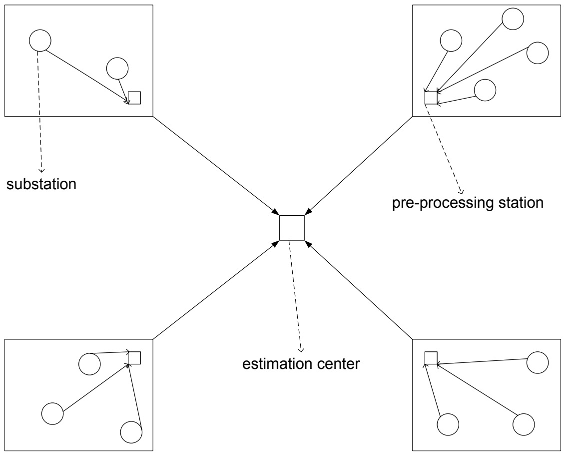

Fig. 1. The Hierarchical System Structure

1. We assume that each substation is metered and each line is metered in bidirections.

Let $\mathbf{x} \in \mathbf{R}^n$ denote the vector composed of the n states of the system, which are the phase angle differences in substations, $\mathbf{z} \in \mathbf{R}^m$ denote the vector composed of the m measurements. With the DC power flow model [12], we have

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{v}, \tag{1}$$

where $\mathbf{H}$ is the m by n measurement matrix with $m > n$. The $\mathbf{v} \in \mathbf{R}^m$ is the random vector composed of the m measurement noises which are independently Gaussian distributed with zero-mean and variance $Var(\mathbf{v})$.

By the DC power flow model, the system states could be estimated using the measurements. One way to reduce the number of measurements transmitted in distance is to use the hierarchical structure as shown in Figure 1. Suppose there is a pre-processing station at each area which combines the measurements inside the area and is responsible for transmitting the processed data to the estimation center. Assume that there is no overlapping measurement, i.e., each measurement is only reported to one of the pre-processing stations. Therefore, let $\mathbf{z}_i$ denote the corresponding measurement of each area, and we have $\mathbf{z} = [\mathbf{z}_1{}' \cdots \mathbf{z}_k{}']'$, $\mathbf{z}_i \in \mathbf{R}^{m_i}$, where m_i is the number of measurements corresponding with area i and $\mathbf{z}'$ denotes the conjugate transpose of $\mathbf{z}$.

We aim to design a distributed two-level linear estimator $\mathbf{K} = \begin{pmatrix} \mathbf{K}_1 & \mathbf{K}_2 & \cdots & \mathbf{K}_k \end{pmatrix}, \mathbf{L} = \begin{pmatrix} \mathbf{L}_1 & \mathbf{L}_2 & \cdots & \mathbf{L}_k \end{pmatrix}$, and $\mathbf{G}_i = \mathbf{L}_i\mathbf{K}_i$, $i = 1, 2, \cdots, k$, where the local measurements related to the pre-processing station i are first locally combined using $\mathbf{K}_i$, and then further processed using $\mathbf{L}_i$ at the estimation center as follows

$$\hat{\mathbf{y}} = \sum_{i=1}^{k} \mathbf{L}_i\mathbf{K}_i\mathbf{z}_i = \sum_{i=1}^{k} \mathbf{G}_i\mathbf{z}_i. \tag{2}$$

Our goal is to design proper $\mathbf{K}$ and $\mathbf{L}$ such that the performance of the proposed distributed estimator is comparable with the MMSE estimator. By assuming that $\mathbf{x}$ and $\mathbf{v}$ are independently Gaussian [13], $\mathbf{z}$ is also Gaussian. The MMSE estimation for $\mathbf{x}$ can be written as follows

$$\hat{\mathbf{x}} = [\mathbf{\Sigma}_x - \mathbf{\Sigma}_x\mathbf{H}'\mathbf{\Sigma}_e^{-1}\mathbf{H}(\mathbf{\Sigma}_x^{-1} + \mathbf{H}'\mathbf{\Sigma}_e^{-1}\mathbf{H})^{-1}]\mathbf{H}'\mathbf{\Sigma}_e^{-1}\mathbf{z}, \tag{3}$$

where $\mathbf{\Sigma}_x$ and $\mathbf{\Sigma}_e$ denote the covariance matrix of $\mathbf{x}$ and $\mathbf{v}$, respectively. Partitioning the corresponding matrices into submatrices for each area, we have $\mathbf{H} = [\mathbf{H}_1{}' \cdots \mathbf{H}_k{}']'$, $\mathbf{H}_i \in \mathbf{R}^{m_i \times n}$, and $\mathbf{\Sigma}_e$ is the block diagonal matrix composed of $\mathbf{\Sigma}_{e_1}, \cdots, \mathbf{\Sigma}_{e_k}$, $\mathbf{\Sigma}_{e_i} \in \mathbf{R}^{m_i \times m_i}$. Let us define

$$\mathbf{W}_i \triangleq [\mathbf{\Sigma}_x - \mathbf{\Sigma}_x\mathbf{H}'\mathbf{\Sigma}_e^{-1}\mathbf{H}(\mathbf{\Sigma}_x^{-1} + \mathbf{H}'\mathbf{\Sigma}_e^{-1}\mathbf{H})^{-1}]\mathbf{H}_i'\mathbf{\Sigma}_{e_i}^{-1} \tag{4}$$

for $i = 1, \cdots, k$ and then (3) can be written as

$$\hat{\mathbf{x}} = \sum_{i=1}^{k} \mathbf{W}_i\mathbf{z}_i. \tag{5}$$

From (2), we can see that reducing the number of measurements transmitted in distance is equivalent to reducing the number of nonzero rows of $\mathbf{K}_i$ because the output of all-zero row of $\mathbf{K}_i$ is always 0 and need not to be transmitted.

Reducing the number of measurements to be transmitted will inevitably cause performance degradation to the distributed estimation. Given the constraints of the number of measurements to be transmitted, the problem of minimizing the gap between the global MMSE estimation $\hat{\mathbf{x}}$ and the distributed estimation $\hat{\mathbf{y}}$, can be formulated as follows,

$$\begin{aligned} &\min_{\mathbf{L}_i, \mathbf{K}_i} E[\|\hat{\mathbf{y}} - \hat{\mathbf{x}}\|_2^2] \\ &\text{s.t.} \quad p_i \leq r_i, \forall i, \end{aligned} \tag{6}$$

where p_i is the number of nonzero rows of $\mathbf{K}_i$, r_i is the communication capability of the pre-processing station i.

III. The Lower Bound For Perfect Measurement Compression

Clearly, if there is no constraint on p_i, $\min_{\mathbf{L}_i, \mathbf{K}_i} E[(\hat{\mathbf{y}} - \hat{\mathbf{x}})^2] = 0$ with $\mathbf{G}_i = \mathbf{W}_i$. In other words, the global MMSE estimation is achieved distributively. In this section, it will be shown that there is a lower bound c_i for r_i. If $r_i \geq c_i$, then it is possible to design appropriate estimators $\mathbf{L}_i, \mathbf{K}_i$ such that the MMSE estimation is achieved distributively. Otherwise, performance degradation of distributed estimation will occur due to the communication constraints.

Lemma 1: For a matrix $\mathbf{A} = \mathbf{B}\mathbf{C}$, where $\mathbf{B}$ and $\mathbf{C}$ are matrices with appropriate dimensions, let f denote the number of nonzero rows of $\mathbf{C}$. Then $f \geq \mathrm{rank}(\mathbf{A})$.

Proof: Suppose that $f < \mathrm{rank}(\mathbf{A})$. Then $\mathrm{rank}(\mathbf{A}) \leq \min(\mathrm{rank}(\mathbf{B}), \mathrm{rank}(\mathbf{C})) \leq f < \mathrm{rank}(\mathbf{A})$. ∎

By Lemma 1, $p_i \geq \mathrm{rank}(\mathbf{G}_i)$, where the equality can be always obtained by the singular value decomposition (SVD).

Thus, $p_i \le r_i$ is equivalent to $\mathrm{rank}(\mathbf{G}_i) \le r_i$. The problem in (6) becomes

$$\min_{\mathbf{L}_i,\mathbf{K}_i} E[\|\hat{\mathbf{y}} - \hat{\mathbf{x}}\|_2^2]$$
$$\text{s.t.} \quad \mathrm{rank}(\mathbf{G}_i) \le r_i, \forall i. \tag{7}$$

Thus, the estimator that we aim to design is actually a low-rank estimator. Performance degradation of distributed estimation may occur due to the constraints of the number of measurements to be transmitted. In the following, we will show that there are some necessary conditions for the communication capabilities such that if they are met, it will be possible to design the low-rank estimators without performance degradation.

Theorem 1: If the measurements cannot be used across the area, a necessary condition to achieve the global MMSE state estimation in the distributed way is $r_i \ge c_i = \mathrm{rank}(\mathbf{H}_i)$.

Proof: The state estimation $\hat{\mathbf{y}}$ in (2) as a function of $\mathbf{z}$, $\hat{\mathbf{y}}(\mathbf{z})$ is said to achieve the MMSE if $\hat{\mathbf{y}}(\mathbf{z}) = \hat{\mathbf{x}}(\mathbf{z})$ for each $\mathbf{z} \in \mathbf{R}^m$. Next, we will show the necessary condition by contradiction.

Suppose there exists an index s such that $r_s < \mathrm{rank}(\mathbf{H}_s)$. It can be clearly seen from (4) that $\mathrm{rank}(\mathbf{W}_s) = \mathrm{rank}(\mathbf{H}_s)$. Since $\mathrm{rank}(\mathbf{G}_s) \le r_s < \mathrm{rank}(\mathbf{H}_s) = \mathrm{rank}(\mathbf{W}_s)$, then $\dim(\Theta_{\mathbf{G}_s}) > \dim(\Theta_{\mathbf{W}_s})$ where $\Theta_{\mathbf{G}_s}$ and $\Theta_{\mathbf{W}_s}$ denote the null space of $\mathbf{G}_s$ and $\mathbf{W}_s$, respectively. In other words, $\exists$ a vector $\tilde{\mathbf{z}}_s$ such that $\mathbf{G}_s\tilde{\mathbf{z}}_s \ne \mathbf{W}_s\tilde{\mathbf{z}}_s$. Therefore, $\exists$ a vector $\tilde{\mathbf{z}} = \begin{pmatrix} \mathbf{0} & \tilde{\mathbf{z}}_s & \mathbf{0} \end{pmatrix}'$ such that $\hat{\mathbf{y}}(\mathbf{z}) = \sum_{i=1}^k \mathbf{G}_i\mathbf{z}_i \ne \hat{\mathbf{x}}(\mathbf{z}) = \sum_{i=1}^k \mathbf{W}_i\mathbf{z}_i$, where $\mathbf{0}$ is the vector with appropriate size composed with all zeros. Obviously, the estimation $\hat{\mathbf{y}}(\mathbf{z})$ does not achieve the MMSE, which leads to a contradiction. Therefore, $r_i \ge \mathrm{rank}(\mathbf{H}_i)$ is a necessary condition for the estimator $\mathbf{G}$ to achieve MMSE. ■

From Theorem 1, we can see that one necessary condition to achieve MMSE is that $r_i \ge \mathrm{rank}(\mathbf{H}_i)$. However, this condition may not be satisfied. In such cases, there exists a tradeoff between the accuracy of the estimation and the number of measurements to be transmitted, which will be shown in next section.

IV. Compression Below the Lower Bound

In this section, we study the scenario where the communication capability is below the lower bound. We first derive an upper bound for the performance degradation, which is the mean square difference between the low-rank estimation and the MMSE estimation, and then design the low-rank estimator by minimizing the upper bound.

Define $d_i \triangleq \sum_{t=1}^i m_t$. The $\mathbf{\Sigma_z}$ is the covariance matrix of $\mathbf{z}$ and is diagonalized by $\mathbf{\Sigma_z} = \mathbf{Q\Lambda_z Q}'$. The $\mathbf{Q}^{ij}$ is the submatrix formed using the elements that appear in rows from d_i to d_{i+1} and columns from d_j to d_{j+1} of matrix $\mathbf{Q}$. The $\mathbf{\Lambda}_{\mathbf{z}^i}$ is the submatrix formed using the elements that appear in rows from d_i to d_{i+1} and columns from d_i to d_{i+1} of matrix $\mathbf{\Lambda_z}$.

Theorem 2: For a system with k pre-processing stations, let $\hat{\mathbf{y}}$ be the low-rank estimation of the state in (7). Then $E[\|\hat{\mathbf{y}} - \hat{\mathbf{x}}\|_2^2] \le k\sum_{j=1}^k\sum_{i=1}^k \left\|\mathbf{F}_i\mathbf{Q}^{ij}\mathbf{\Lambda}_{\mathbf{z}^j}^{1/2}\right\|_F^2$ where $\mathbf{F}_i \triangleq \mathbf{G}_i - \mathbf{W}_i$ and $\mathbf{F} \triangleq \begin{pmatrix} \mathbf{F}_1 & \mathbf{F}_2 & \cdots & \mathbf{F}_k \end{pmatrix}$.

Proof: The performance degradation can be expressed by

$$E[\|\hat{\mathbf{y}} - \hat{\mathbf{x}}\|_2^2] = \mathrm{tr}(E[(\hat{\mathbf{y}} - \hat{\mathbf{x}})(\hat{\mathbf{y}} - \hat{\mathbf{x}})']) = \mathrm{tr}(\mathbf{F\Sigma_z F}')$$

Note that the $\mathbf{\Sigma_z} = \mathbf{H\Sigma_x H}' + \mathbf{\Sigma_v}$, which is positive definite and thus can be diagonalized by $\mathbf{\Sigma_z} = \mathbf{Q\Lambda_z Q}'$. Then, $\mathrm{tr}(\mathbf{F\Sigma_z F}') = \mathrm{tr}(\mathbf{FQ\Lambda_z Q}'\mathbf{F}') = \left\|\mathbf{FQ\Lambda_z}^{1/2}\right\|_F^2$. Substituting the submatrices into it, we have

$$\begin{aligned}
& \left\|\mathbf{FQ\Lambda_z}^{1/2}\right\|_F^2 \\
= & \Big\|\Big(\textstyle\sum_{i=1}^k \mathbf{F}_i\mathbf{Q}^{i1}\mathbf{\Lambda}_{\mathbf{z}^1}^{1/2} \quad \sum_{i=1}^k \mathbf{F}_i\mathbf{Q}^{i2}\mathbf{\Lambda}_{\mathbf{z}^2}^{1/2} \\
& \quad \cdots \quad \textstyle\sum_{i=1}^k \mathbf{F}_i\mathbf{Q}^{ik}\mathbf{\Lambda}_{\mathbf{z}^k}^{1/2} \Big)\Big\|_F^2 \\
= & \sum_{j=1}^k \left\| \left(\sum_{i=1}^k \mathbf{F}_i\mathbf{Q}^{ij}\mathbf{\Lambda}_{\mathbf{z}^j}^{1/2} \right) \right\|_F^2 \\
\le & \ k\sum_{i=1}^k\sum_{j=1}^k \left\|\mathbf{F}_i\mathbf{Q}^{ij}\mathbf{\Lambda}_{\mathbf{z}^j}^{1/2}\right\|_F^2,
\end{aligned} \tag{8}$$

where (8) is due to the Cauchy-Schwartz inequality. ■

Define $\mathbf{D}_i \triangleq \begin{pmatrix} \mathbf{Q}^{i1}\mathbf{\Lambda}_{\mathbf{z}^1}^{1/2} & \cdots & \mathbf{Q}^{ik}\mathbf{\Lambda}_{\mathbf{z}^k}^{1/2} \end{pmatrix} \in \mathbf{R}^{m_i \times m}$. Since $\sum_{j=1}^k \left\|\mathbf{F}_i\mathbf{Q}^{ij}\mathbf{\Lambda}_{\mathbf{z}^j}^{1/2}\right\|_F^2 = \|\mathbf{F}_i\mathbf{D}_i\|_F^2$, by minimizing the upper bound of the performance degradation, (7) becomes

$$\min_{\mathbf{F}_i} \sum_{i=1}^k \|\mathbf{F}_i\mathbf{D}_i\|_F^2$$
$$s.t. \quad \mathrm{rank}(\mathbf{G}_i) \le r_i, \forall i \tag{9}$$

Applying SVD to $\mathbf{D}_i$, $\mathbf{D}_i = \mathbf{U}_i\mathbf{\Sigma}_i\mathbf{V}_i'$. Since $\mathbf{V}_i$ is a unitary matrix, $\|\mathbf{F}_i\mathbf{D}_i\|_F^2 = \|\mathbf{F}_i\mathbf{U}_i\mathbf{\Sigma}_i\mathbf{V}_i'\|_F^2 = \|\mathbf{F}_i\mathbf{U}_i\mathbf{\Sigma}_i\|_F^2 = \left\| \begin{matrix} \mathbf{F}_i\mathbf{U}_i\mathbf{\Sigma}_i^* & \mathbf{0} \end{matrix} \right\|_F^2 = \|\mathbf{F}_i\mathbf{U}_i\mathbf{\Sigma}_i^*\|_F^2$ where $\mathbf{\Sigma}_i = \begin{pmatrix} \mathbf{\Sigma}_i^* & \mathbf{0} \end{pmatrix}$, $\mathbf{\Sigma}_i^* \in \mathbf{R}^{m_i \times m_i}$. Therefore,

$$\|\mathbf{F}_i\mathbf{D}_i\|_F^2 = \|(\mathbf{G}_i - \mathbf{W}_i)\mathbf{U}_i\mathbf{\Sigma}_i^*\|_F^2 = \|(\mathbf{G}_i\mathbf{U}_i - \mathbf{W}_i\mathbf{U}_i)\mathbf{\Sigma}_i^*\|_F^2. \tag{10}$$

Thus, the estimator $\mathbf{G}_i$ is

$$\mathbf{G}_i = \arg\min_{\mathbf{G}_i} \|(\mathbf{G}_i\mathbf{U}_i - \mathbf{W}_i\mathbf{U}_i)\mathbf{\Sigma}_i^*\|_F^2$$
$$s.t. \quad \mathrm{rank}(\mathbf{G}_i) \le r_i, \forall i, \tag{11}$$

which can be transformed into the weighted low rank approximation problem and solved by the numerical method in [14].

V. Simulation Results

In this section, we evaluate the proposed scheme under the IEEE 14 bus system using Matpower [15]. We divide the 14 substations into two groups. The measurements are first processed in the pre-processing station and then transmitted to the estimation center for state estimation using proposed scheme. The performance of the estimators is evaluated in terms of mean square error as follows,

$$\text{MMSE} = \frac{\mathbf{\Sigma}_{i=1}^N\|\hat{\mathbf{x}} - \mathbf{x}\|_2^2}{N}$$
$$\text{LRMSE} = \frac{\mathbf{\Sigma}_{i=1}^N\|\hat{\mathbf{y}} - \mathbf{x}\|_2^2}{N},$$

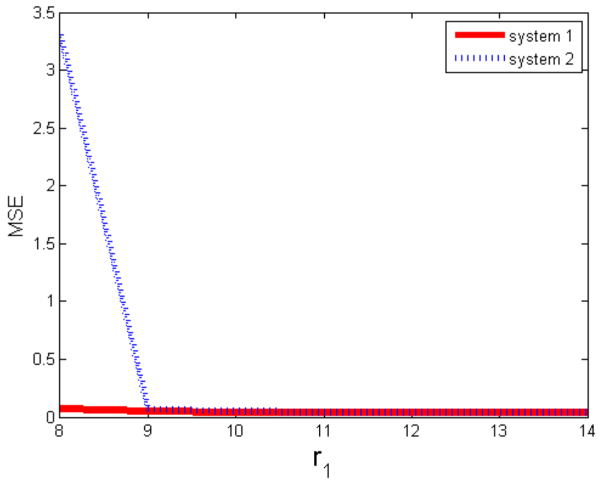

Fig. 2. The Comparison of Two Systems of Different Topology

where N is the number of runs, MMSE is the minimum mean square error and LRMSE is the mean square error achieved by the distributed low-rank estimator. Without loss of generality, the 13 state components and the 54 measurement noise components are assumed to be independent zero-mean Gaussian variables with variance 4 for the state components and 0.2 for the measurement noise components, respectively.

TABLE I
THE MSE FOR $\mathrm{rank}(\mathbf{H}_1) = 10, \mathrm{rank}(\mathbf{H}_2) = 12$

r_1	r_2	MMSE	LRMSE	r_1	r_2	MMSE	LRMSE
8	10	0.04035	0.44954	10	10	0.03917	0.07109
8	11	0.03942	0.41249	10	11	0.03992	0.04627
8	12	0.03909	0.40263	10	12	0.03943	0.03943
8	13	0.03935	0.39911	10	13	0.03875	0.03875
9	10	0.03915	0.07451	11	10	0.03933	0.07145
9	11	0.03997	0.04889	11	11	0.03973	0.04597
9	12	0.03926	0.04153	11	12	0.03978	0.03978
9	13	0.03948	0.04167	11	13	0.03971	0.03971

In Table I, it is illustrated that if $r_1 \geq \mathrm{rank}(\mathbf{H}_1)$, $r_2 \geq \mathrm{rank}(\mathbf{H}_2)$, then the low-rank estimator achieves the MMSE. Otherwise, the LRMSE is higher than the MMSE. The smaller the communication capability is, the higher is the LRMSE.

By changing the grouping of these substations, it can be shown in Figure 2 that the two systems have quite different performances even with the same communication settings. In system 1, the substations $1, 2, 3, 4, 5, 6, 11$ are in one group while the rest are in the other group. In system 2, the substations $1, 2, 3, 4, 5, 6, 11, 14$ are in one group while the rest are in the other group. Fixing $r_2 = 9$ which is the lower bound for r_2 of both systems and vary r_1, the performances of the two systems have distinctions at some r_1. It illustrates that the lower bounds for r_1, r_2 are dependent on the grouping, i.e., the topology of the system.

VI. CONCLUSION

In this paper, we present a solution to the distributed state estimation problem in a smart grid with communication constraints by designing a low-rank estimator. It is shown that to achieve the global MMSE estimation, there exists a minimum requirement of the communication capability. If the requirement is satisfied, a distributed estimator can be designed to achieve the global MMSE estimation. Otherwise, an estimator meeting such constraints is proposed to trade the performance for communication cost.

A number of issues can be further addressed in this paper. For instance, since directly solving the problem in (7) is rather difficult, we tackle the optimization problem by minimizing the upper bound of the performance degradation derived in (8). Such an approximation can be further improved by either finding a tighter upper bound of the performance degradation or directly solving the original optimization problem if possible. Moreover, since the states of a system are actually the phase difference in different substations according to the DC power flow model, instead of Gaussian distribution, Laplace distribution may be an alternative model for the system states.

REFERENCES

[1] Y. Liu, P. Ning, and M. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security (TISSEC)*, vol. 14, no. 1, p. 13, 2011.

[2] M. Saif and Y. Guan, "Decentralized state estimation in large-scale interconnected dynamical systems," *Automatica*, vol. 28, no. 1, pp. 215 – 219, 1992.

[3] U. Khan and J. Moura, "Distributing the kalman filter for large-scale systems," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4919 –4935, oct. 2008.

[4] S. Stankovic, M. Stankovic, and D. Stipanovic, "Consensus based overlapping decentralized estimator," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, pp. 410 –415, feb. 2009.

[5] "Wolfe island wind project (197.8mw)." [Online]. Available: http://www.powerauthority.on.ca/wind-power/wolfe-island-wind-project-1978-mw-wolfe-island

[6] G. Valverde and V. Terzija, "Unscented kalman filter for power system dynamic state estimation," *Generation, Transmission Distribution, IET*, vol. 5, no. 1, pp. 29 –37, jan. 2011.

[7] T. Van Cutsem, J. Horward, and M. Ribbens-Pavella, "A two-level static state estimator for electric power systems," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-100, no. 8, pp. 3722 –3732, aug. 1981.

[8] A. Gomez-Exposito and A. de la Villa Jaen, "Two-level state estimation with local measurement pre-processing," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 676 –684, may 2009.

[9] U. Khan, M. Ilic, and J. Moura, "Cooperation for aggregating complex electric power networks to ensure system observability," in *2008 First International Conference on Infrastructure Systems and Services: Building Networks for a Brighter Future (INFRA)*, nov. 2008, pp. 1 –6.

[10] S. Kar, J. Moura, and K. Ramanan, "Distributed parameter estimation in sensor networks: Nonlinear observation models and imperfect communication," *IEEE Transactions on Information Theory*, vol. 58, no. 6, pp. 3575 –3605, june 2012.

[11] L. Xie, D.-H. Choi, S. Kar, and H. V. Poor, "Fully distributed state estimation for wide-area monitoring systems," *IEEE Transactions on Smart Grid*, vol. PP, no. 99, p. 1, 2012.

[12] A. Abur and A. Exposito, *Power system state estimation: theory and implementation*. CRC, 2004, vol. 24.

[13] O. Kosut, L. Jia, R. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 645 –658, dec. 2011.

[14] N. Srebro and T. Jaakkola, "Weighted low-rank approximations," in *Proceedings of the Twentieth International Conference on Machine Learning*, vol. 20, no. 2, 2003, p. 720.

[15] R. Zimmerman, C. Murillo-Sanchez, and R. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12 –19, feb. 2011.