State Estimation with Sampling Offsets in Wide Area Measurement Systems

Hoi-To Wai and Anna Scaglione

Department of Electrical and Computer Engineering, University of California at Davis, Davis, CA 95616, USA
Email: {htwai,ascaglione}@ucdavis.edu

***Abstract*—An implicit assumption made in studies on state estimation is that the time and frequency at which these measurements are taken is consistent across all the distributed sensors. For instance, in the literatures on Wide Area Measurement Systems (WAMS) deployed in the power grid, where the sensors equipped with Global Positioning Signals (GPS), the sensors are deemed capable to provide perfectly synchronous readings at the various sampling sites. The validity of the assumption may need to be re-examined with the recent advancements in decentralized state estimation algorithms. Importantly, when there are timing offsets between sampling devices, the effects on the measurement system's performance can be catastrophic. The prevalent point of view is to either study the resulting error, or to resort to Kalman filtering for aligning the measurements. Taking on this view typically requires additional information about the underlying state. In this paper, we revisit the problem of state estimation and propose a new model for data acquisition under asynchronous sampling. The key idea is to apply sampling theory and to exploit the redundancy in the spatial sampling to interpolate the system state. We provide a necessary and sufficient condition for identifiability of the time offsets and propose an algorithm for the joint regression on state and timing offsets. The efficacy of the proposed algorithm is shown by numerical simulations.**

***Index Terms*—state estimation, sampling offsets, smart grid**

I. Introduction

The aim of sensor networks in cyber-physical systems is to provide measurements that bear information on the state of the physical infrastructure for monitoring and control. For instance, in interconnected power grids, Wide Area Measurement Systems (WAMS) are emerging as the new sensing infrastructures to provide situational awareness in real-time on what is the grid state [1], [2].

However, one of the implicit assumption in these papers is that the measurements are obtained at the sensors in a perfectly synchronous manner. This is because, either the system state evolution is sufficiently slow that the lack of synchrony in sampling is negligible, or the timing information is sufficiently accurate to realign the measurements as if they were taken with a consistent time reference prior to processing. Another plausible violation to these assumptions occurs when the sensor's clocks are tampered by malicious attackers, and therefore timing can no longer be trusted [3]. It is important to employ a sensor fusion algorithm that is resilient to timing error.

Most of the prior arts concerned with mitigating timing errors are based on Kalman Filtering [4], [5]. There, the unknown timing error is jointly estimated as a fixed parameter with the time varying system state. However, it is assumed that the dynamic equations that represent the evolution of system state is given a priori. As such, these algorithms are usually developed only for specific applications in which this information is either available or easy to estimate. Another related issue to ours is that of the blind calibration problem for time-interleaved ADCs [6], [7]. Though these works are aimed at a different application, their goal is similar to ours, i.e., to estimate the unknown asynchrony between the ADCs that sample at interleaved time, in order to removing aliasing in each of the interleaved ADCs.

In this paper, we take the view that if the relationship between measurements and state is memoryless, the regression problem and sampling theorem can be used in combination to determine both the state and timing errors. This amounts to using a sampling expansion of the system state as a formal way to capture its *smoothness* and as an alternative to assuming a dynamical model. Specifically, we model the system states as a band-limited continuous-time signal and the measurements as samples taken at the critical sampling frequency of the system states. As a consequence, the relationship between the system states and the measurements (under timing error) can be expressed in frequency domain using discrete time Fourier transform (DTFT). Using the DTFT representation, we derive a necessary and sufficient condition to guarantee unique recovery of state and time offsets in the noiseless case. The joint regression problem is then cast as a least square problem on both the system state (i.e., a series of coefficient vectors) and timing error. Based upon this, we develop an algorithm for tackling the joint regression problem. We show that the algorithm can be implemented in a decentralized manner using Gossip-based exchanges. Our simulation results demonstrate that the proposed algorithm (decentralized and centralized) can reliably recover the system states under asynchronous sampling.

Notations: We follow the standard mathematical notations used in signal processing literature. For instance, $\mathcal{N}(\mathbf{A})$ denotes the nullspace of the matrix $\mathbf{A}$ such that $\mathbf{y} \in \mathcal{N}(\mathbf{A})$ if $\mathbf{A}\mathbf{y} = \mathbf{0}$; $\dim(\mathcal{V})$ denotes the dimension of a vector space $\mathcal{V}$.

II. System model

Consider a band-limited continuous-time state vector $\mathbf{x}_c(t) \in \mathbb{C}^N$ with bandwidth $W/2$ Hz. Let there be K sensors in the measurement system which measure $\mathbf{x}_c(t)$ according to the following model. In particular, at time t, sensor i has access to a measured signal that is linearly related to the state vector:

$$\boldsymbol{\zeta}_i(t) = \mathbf{H}_i \mathbf{x}_c(t) + \mathbf{w}_i(t), \tag{1}$$

where $\mathbf{H}_i \in \mathbb{C}^{M_i \times N}$ is the measurement matrix. For example, in the grid, the signal $\mathbf{x}_c(t)$ in (1) are bus voltages and $\boldsymbol{\zeta}_i(t)$ are measured currents/voltages. For ease of exposition, we assume that $M_i = M$.

The relationship between $\mathbf{x}_c(t)$ and $\boldsymbol{\zeta}_i(t)$ is memoryless and linear. Hence, the observed signal $\boldsymbol{\zeta}_i(t)$ is also band limited with bandwidth $W/2$ Hz. Now, consider the case with critical sampling, the K sensors measurement streams are sampled every $T_s = 1/W$ second but with a random unknown offset β_i. The nth sample from sensor i are taken at time t_n^i (relative to an absolute time scale t):

$$\frac{n}{W} = t_n^i + \beta_i \quad \rightarrow \quad t_n^i = \frac{n}{W} - \beta_i. \tag{2}$$

By denoting:

$$\mathbf{x}(n) = \mathbf{x}_c(nT_s), \quad b_i = W\beta_i, \quad f(n - b_i) = \frac{\sin(\pi(n-b_i))}{\pi(n-b_i)}.$$

This material is based upon work supported by the Department of Energy under Award Number DE-OE0000097.

The Whittaker-Shannon interpolation formula implies that the nth sample from sensor i can be expressed as

$$\begin{aligned} \boldsymbol{\zeta}_i(n) &\triangleq \boldsymbol{\zeta}_i(t_n^i) = \mathbf{H}_i[\mathbf{x}(n) \star f(n-b_i)] + \mathbf{w}_i(t_n^i) \quad (3) \\ &= \mathbf{H}_i \sum_{k=-\infty}^{\infty} \mathbf{x}(k) f(n-k-b_i) + \mathbf{w}_i(n). \quad (4) \end{aligned}$$

As the DTFT of $f(n-b_i)$ is given by $F(e^{j\omega}) = e^{-j\omega b_i}$, thus, applying discrete time Fourier transform (DTFT) to the both sides of (12) gives:

$$\mathbf{Z}_i(e^{j\omega}) = \mathbf{H}_i \mathbf{X}(e^{j\omega}) e^{-j\omega b_i} + \mathbf{W}_i(e^{j\omega}). \quad (5)$$

Our task, in the presence of sampling offsets, is to estimate the spectrum $\mathbf{X}(e^{j\omega})$ *without* knowing the offset error b_i. Note that changing all b_i by a constant c leads to an identical representation, albeit the estimated spectrum will be $\mathbf{X}(e^{j\omega})e^{j\omega c}$. To avoid such ambiguity, we constrain one of the sensors to be a reference node. In particular, we assume

$$\beta_1 = 0, \quad |W\beta_i| < B. \quad (6)$$

Under the assumption that the noise is i.i.d. Gaussian, it is reasonable for us to solve a Non Linear Least Square (NLLS) regression problem to accomplish the task. However, it is also well known that the NLLS regression problem may have multiple (*both* global and local) optimal solutions and there is only one optimal solution that gives the true state estimate. Hence, before delving into the computational aspect of the NLLS regression problem, we shall investigate under what condition can observability be guaranteed.

A. Identifiability condition

We consider (5) under noiseless observations, i.e., by setting $\mathbf{W}_i(e^{j\omega}) = \mathbf{0}$. Let $\mathbf{b} = (b_2, \ldots, b_K)$ and suppose that $(\mathbf{X}(e^{j\omega}), \mathbf{b})$ is the true state and set of sampling offsets. The noiseless observation model in frequency domain is:

$$\begin{aligned} &\mathbf{Z}_1(e^{j\omega}) = \mathbf{H}_1 \mathbf{X}(e^{j\omega}), \quad (7) \\ &\mathbf{Z}_i(e^{j\omega}) = \mathbf{H}_i \mathbf{X}(e^{j\omega}) e^{-j\omega b_i}, \qquad i = 2, ..., K. \end{aligned}$$

Let $\epsilon_i \triangleq \hat{b}_i - b_i$ and $\boldsymbol{\epsilon} \triangleq (\epsilon_2, ..., \epsilon_K)$. Any solution to (7), $(\hat{\mathbf{X}}(e^{j\omega}), \hat{\mathbf{b}})$, must satisfy

$$\begin{aligned} &\mathbf{0} = \mathbf{H}_1[\mathbf{X}(e^{j\omega}) - \hat{\mathbf{X}}(e^{j\omega})], \quad (8) \\ &\mathbf{0} = \mathbf{H}_i[\mathbf{X}(e^{j\omega}) - \hat{\mathbf{X}}(e^{j\omega}) e^{-j\omega\epsilon_i}], \qquad i = 2, ..., K. \quad (9) \end{aligned}$$

This leads to the following necessary and sufficient condition:

Lemma 1. *Assume* $\mathbf{X}(e^{j\omega}) \neq 0$ *and consider:*

$$\boldsymbol{\mathcal{H}}(\boldsymbol{\epsilon}) \triangleq \begin{bmatrix} \mathbf{H}_1 & -\mathbf{H}_1 \\ \mathbf{H}_2 & -e^{-j\omega\epsilon_2}\mathbf{H}_2 \\ \vdots & \vdots \\ \mathbf{H}_K & -e^{-j\omega\epsilon_K}\mathbf{H}_K \end{bmatrix}.$$

The following two statements are equivalent:
1. $\boldsymbol{\mathcal{H}}(\boldsymbol{\epsilon})$ *is full column rank when* $\epsilon_i \neq 0$ *for some* i
2. $(\mathbf{X}(e^{j\omega}), \mathbf{b})$ *is the unique tuple that satisfies* (7).

Proof. Equation (8)–(9) is equivalent to

$$\boldsymbol{\mathcal{H}}(\boldsymbol{\epsilon}) \begin{bmatrix} \mathbf{X}(e^{j\omega}) \\ \hat{\mathbf{X}}(e^{j\omega}) \end{bmatrix} = \mathbf{0} \quad (10)$$

We see that $\hat{\mathbf{b}} = \mathbf{b}$ if Statement 1 is true since the system of equations has no other solution than $\mathbf{0}$ (which is impossible as $\mathbf{X}(e^{j\omega}) \neq \mathbf{0}$) whenever $\hat{\mathbf{b}} \neq \mathbf{b}$. This, together with (7), implies that $(\hat{\mathbf{X}}(e^{j\omega}), \mathbf{b}) \equiv (\mathbf{X}(e^{j\omega}), \mathbf{b})$ as the tall matrix $[\mathbf{H}_1^T \ \cdots \ \mathbf{H}_K^T]^T$ is full rank. On the other hand, suppose the contrary to Statement 1 such that $\boldsymbol{\mathcal{H}}(\boldsymbol{\epsilon})$ is not full column rank for all $\boldsymbol{\epsilon} \in \mathbb{R}^K$. There exists $(\hat{\mathbf{X}}(e^{j\omega}), \hat{\mathbf{b}}) \neq (\mathbf{X}(e^{j\omega}), \mathbf{b})$ that satisfies (10), i.e., $(\mathbf{X}(e^{j\omega}), \mathbf{b})$ is not unique. This shows that Statement 1 is a sufficient and necessary condition for identifiability. **Q.E.D.**

Lemma 1 seems to indicate that the system observability is quite fragile to timing errors, and that the demand for number of sensors doubles when such errors are possible for all sensors.

To understand the sufficient condition better, one can assume that a subset of the sensors are synchronized, but there are at most $K' < K$ outliers, like for instance in the case of a Byzantine attack. Suppose that the attackers have not attacked the reference node and that the last K' sensors are attacked. We can group all the sensors with a coherent clock and treat them as a single sensor:

$$\mathbf{Z}'_1(e^{j\omega}) = \begin{bmatrix} \mathbf{Z}_1(e^{j\omega}) \\ \vdots \\ \mathbf{Z}_{K-K'-1}(e^{j\omega}) \end{bmatrix}, \quad \mathbf{H}'_1 = \begin{bmatrix} \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_{K-K'-1} \end{bmatrix}.$$

The rest of the sensors are attacked such that their clocks are offsetted by $\mathbf{b}_{K'} = (b_{K'}, \ldots, b_K)$. The observations at these nodes are:

$$\mathbf{Z}'_2(e^{j\omega}) = \begin{bmatrix} \mathbf{Z}_{K'}(e^{j\omega}) \\ \vdots \\ \mathbf{Z}_K(e^{j\omega}) \end{bmatrix}, \quad \mathbf{H}'_2(\mathbf{b}_{K'}) = \begin{bmatrix} e^{j\omega b_{K'}}\mathbf{H}_{K'} \\ \vdots \\ e^{j\omega b_K}\mathbf{H}_K \end{bmatrix}$$

Using linear algebra arguments, it is clear that the attack cannot be successful if:

$$\mathcal{N}(\mathbf{H}'_1) \cap \mathcal{N}(\mathbf{H}'_2(\mathbf{b}_{K'})) = \{\mathbf{0}\}, \ \forall \ \mathbf{b}_{K'} \neq \mathbf{0} \quad (11)$$

On the other hand, when (11) holds, $\boldsymbol{\mathcal{H}}(\boldsymbol{\epsilon})$ is always a full column rank matrix if $\epsilon_i \neq 0$ for some i. Thus, applying Lemma 1 reveals that the attack vector $\mathbf{b}$ can be estimated uniquely and the attack fails.

Finally, the following corollary can be obtained from Lemma 1.

Corollary 1. $(\mathbf{X}(e^{j\omega}), \mathbf{b})$ *is uniquely identifiable only if* $KM \geq 2N$ *and if for* $\mathbf{b} = 0$ *the state is observable with half of the sensors.*

III. Proposed Algorithm

This section describes an algorithm for tackling the state estimation problem in the presence of asynchronous sampling. We notice that to estimate the timing errors, it is essential to exploit the temporal relationship between measured samples. Ideally, we should obtain the DTFTs from an infinite number of samples. However, doing so is impossible due to the memory limitation in practical systems. To obtain a trade-off between accuracy and computational complexity, our strategy is to process the asynchronous samples in a block-based manner. To this end, we notice that as the function $f(n-b_i) = O(n^{-2})$ for non-integer b_i, the samples of $f(n-b_i)$ vanishes for $|n-b_i| > J$ with large J. Now, by retaining only the samples for which $|n-b_i| \leq J$, with $P \triangleq J + B$, we have

$$\boldsymbol{\zeta}_i(n) \approx \mathbf{H}_i \sum_{k=n-P}^{n+P} \mathbf{x}(k) f(n-k-b_i) + \mathbf{w}_i(n). \quad (12)$$

Therefore, by observing $L+1 \geq 2P+1$ samples, each sensor will be able to obtain a reasonably good estimate of $\hat{\mathbf{Z}}_i(e^{j\omega}) \approx \mathbf{Z}_i(e^{j\omega})$. Now, by taking a possibly weighted discrete Fourier transform (DFT) of $L+1$ samples, the NLLS regression problem for joint state and timing estimation is given as:

$$\begin{aligned} \min_{\mathbf{X}(e^{j\omega}),\mathbf{b}} \quad & \sum_{i=1}^{K}\sum_{\omega\in\Omega} \|\mathbf{Z}_i(e^{j\omega}) - e^{-j\omega b_i}\mathbf{H}_i\mathbf{X}(e^{j\omega})\|_2^2 \\ \text{s.t.} \quad & b_1 = 0,\ \ |b_i| < B,\ i = 2,...,K, \end{aligned} \tag{13}$$

where $\Omega = \{2\pi\ell/L\}_{\ell=-L/2}^{L/2}$ and $|\Omega| = L+1$.

Problem (13) is non-convex due to its dependence on both $\mathbf{X}(e^{j\omega})$ and b_i. A natural approach for tackling it is to apply an alternating optimization (AO) algorithm. In particular, let k denotes the AO iteration index, we alternate between the update of $\mathbf{b}^{(k)}$ and that of $\mathbf{X}^{(k)}(e^{j\omega})$ iteratively as follows, i.e., let $\boldsymbol{H}^\dagger$ be the pseudo-inverse of matrix $\boldsymbol{H}$ and

$$\boldsymbol{H} = \begin{bmatrix} \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_K \end{bmatrix},\ \mathbf{Z}^{(k)}(e^{j\omega}) = \begin{bmatrix} \mathbf{Z}_1(e^{j\omega}) \\ e^{j\omega b_2^{(k)}}\mathbf{Z}_2(e^{j\omega}) \vdots \\ e^{j\omega b_K^{(k)}}\mathbf{Z}_K(e^{j\omega}) \end{bmatrix}. \tag{14}$$

The AO procedure is given by the following update equations:

$$\mathbf{X}^{(k+1)}(e^{j\omega}) = \boldsymbol{H}^\dagger\mathbf{Z}^{(k)}(e^{j\omega}) \tag{15}$$

$$b_i^{(k+1)} = \mathcal{P}_{\mathcal{B}}\left(b_i^{(k)} + \rho_k \sum_{\omega\in\Omega} a_i(\omega)\sin(f_i(\omega) - \omega b_i^{(k)})\omega\right), \tag{16}$$

where $\mathcal{B} = [-B, B]$ is the interval that constrains b_i, $\rho_k > 0$ is a step size to be found using a line search procedure, e.g., the Armijo line search [8], and $a_i(\omega), f_i(\omega) \in \mathbb{R}$ are given by

$$\mathbf{Z}_i^H(e^{j\omega})\mathbf{H}_i\mathbf{X}^{(k+1)}(e^{j\omega}) \equiv a_i(\omega)e^{jf_i(\omega)}.$$

The latter term in (15) is the projected gradient taken at $b_i^{(k)}$ with respect to b_i. To attain faster convergence, wiithin one AO iteration, (16) might be repeated for several times. The resultant algorithm falls into the class of *inexact* block coordinate descent which is recently analyzed in [9]. In particular, the AO steps (15)–(16) converges to a local minimum of (13) as $k \to \infty$.

A. Decentralized implementation using Gossip-based exchange

We now describe a decentralized implementation strategy for the proposed algorithm. Notice that a decentralized algorithm is often preferable for WAMS as the fusion center may be absent. In (15)–(16), the b_i's update (16) is clearly decoupled with the updates for other sensors. Hence, our only task is to tackle the update (15) in a completely decentralized manner. To this end, existing decentralized algorithms such as gossip-based update [1], diffusion LMS [10], constrained consensus [11], etc., may be applied. Among them, the gossip-based update is relevant as its convergence rate is faster.

To describe the algorithm, we first see that the update (15) can be expressed as:

$$\mathbf{X}^{(k+1)}(e^{j\omega}) = \Big(\sum_{i=1}^{K}\mathbf{H}_i^H\mathbf{H}_i\Big)^{-1}\Big(\sum_{i=1}^{K}\mathbf{H}_i^H\mathbf{Z}_i^{(k)}(e^{j\omega})\Big) \tag{17}$$

Assuming that the sub-matrix $\mathbf{H}_i$ is known to the ith sensor and noting that $\mathbf{Z}_i^{(k)}(e^{j\omega})$ is also known to the ith sensor. Evaluating (17) merely involves computing the average terms inside the two brackets in (17). Following the idea from [1], these terms can be computed distributively using Gossip exchanges. In particular, let m denotes the Gossip iteration index and we define:

$$\boldsymbol{h}_i^{(0)} = \mathbf{H}_i^H\mathbf{H}_i,\ \boldsymbol{z}_i^{(k,0)}(e^{j\omega}) = \mathbf{H}_i^H\mathbf{Z}_i^{(k)}(e^{j\omega})$$

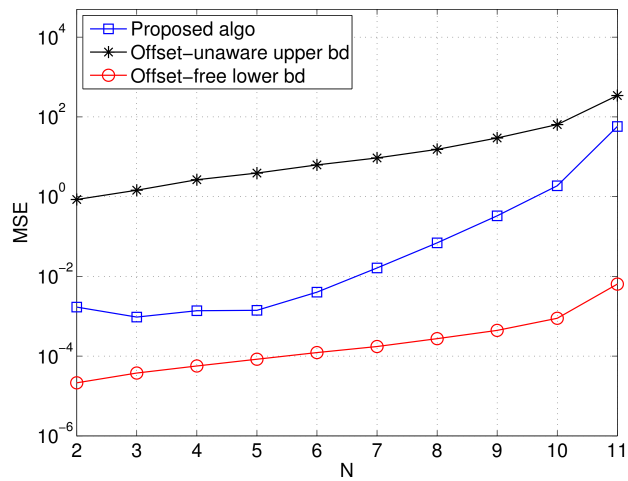

Fig. 1. The average MSE of $\mathbf{X}(e^{j\omega})$ against dimension of the state N. $M = 3, K = 4, L = 41$.

The Gossip exchange iterations for average consensus is given by:

$$\boldsymbol{h}_i^{(m+1)} = \textstyle\sum_{j=1}^{K} W_{ij}^{(m+1)}\boldsymbol{h}_j^{(m)}, \tag{18}$$

$$\boldsymbol{z}_i^{(k,m+1)}(e^{j\omega}) = \textstyle\sum_{j=1}^{K} W_{ij}^{(m+1)}\boldsymbol{z}_j^{(k,m)}(e^{j\omega}), \tag{19}$$

where $W_{ij}^{(m+1)}$ is the (i,j)th entry of a doubly stochastic matrix $\mathbf{W}^{(m+1)}$. Employing the uncoordinated random exchange (URE) protocol [1], at iteration m, a sensor i_m wakes up and randomly selects one of its neighbors $j_m \in \mathcal{N}_{j_m}$. The two sensors i_m and j_m exchanges information and $\mathbf{W}^{(m)} = 0.5(\mathbf{e}_{i_m} + \mathbf{e}_{j_m})(\mathbf{e}_{i_m} + \mathbf{e}_{j_m})^T$. Repeating the same process and assuming that the underlying graph between sensors is connected, it is guaranteed that as $m \to \infty$:

$$\boldsymbol{h}_i^{(m)} = \frac{1}{K}\sum_{i=1}^{K}\mathbf{H}_i^H\mathbf{H}_i,\ \boldsymbol{z}_i^{(k,m)}(e^{j\omega}) = \frac{1}{K}\sum_{i=1}^{K}\mathbf{H}_i^H\mathbf{Z}_i^{(k)}(e^{j\omega}),$$

for all i. This is exactly what we want in (17). Therefore, (17) can be replaced by the above decentralized computation and we compute $\hat{\mathbf{X}}_i^{(k+1)}(e^{j\omega})$ at the ith sensor using $\boldsymbol{h}_i^{(m)}$ and $\boldsymbol{z}_i^{(k,m)}(e^{j\omega})$. We remark that the average consensus for $\boldsymbol{h}_i^{(m)}$ needs not be repeated for each of the kth AO iteration.

IV. Numerical Results & Conclusion

This section investigates the effects of sampling offsets on state estimation using numerical examples. The first part of this section considers the case when the state and measurement are generated using synthetic data. The spectrum $\mathbf{X}(e^{j\omega})$ is generated by performing an L-point DFT on L complex zero-mean unit variance white Gaussian generated samples, while the spectrum of the asynchronous measurements $\mathbf{Z}_i(e^{j\omega})$ is obtained according to (5). The additive noise $\mathbf{V}_i(e^{j\omega})$ is a complex white Gaussian random vector with variance of $\sigma_w = 10^{-2}$. We also assume $W = 1$ and the offset error is uniformly distributed over $[-0.2, 0.2]$. For the simulation results, we focus on comparing the mean square error (MSE) in estimating the DFT spectrum, i.e., the average of $\frac{1}{L}\sum_{\omega\in\Omega}\|\hat{\mathbf{X}}(e^{j\omega}) - \mathbf{X}(e^{j\omega})\|_2^2$ where $\hat{\mathbf{X}}(e^{j\omega})$ is the estimated spectrum. The performance of the proposed method is compared to two benchmarks: the curve 'offset-free lower bd' refers to the case when no offset error is introduced; the curve 'offset-unware bound', instead, refers to the case when the state estimator is unaware of the offset error. For the AO algorithm, we initialize it by setting $b_i = 0$ for all i and terminate it with

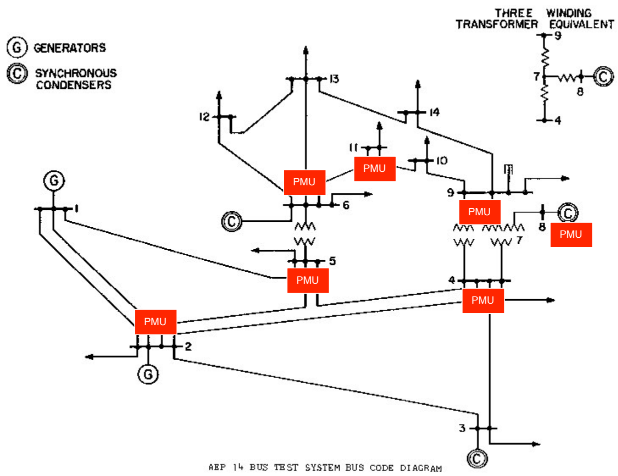

Fig. 2. The IEEE-14 bus system. The locations of the PMU placements are indicated by red. The PMUs are placed on bus $\{2, 4, 5, 6, 8, 9, 11\}$.

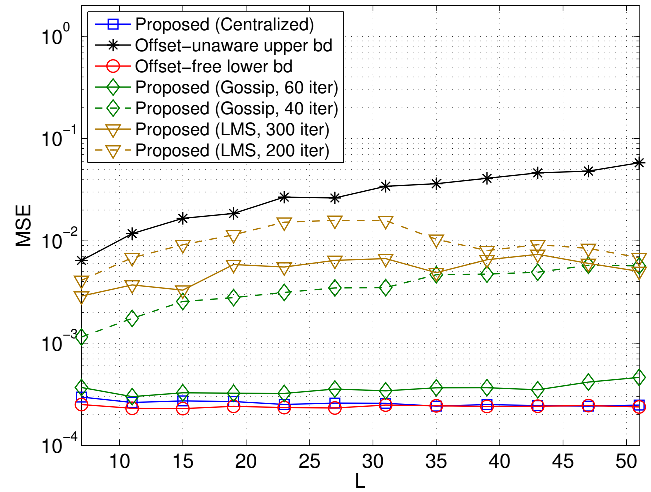

Fig. 3. The average MSE of $\mathbf{X}(e^{j\omega})$ against the frame size L using the IEEE-14 bus system in Fig. 2.

$\max\{L, 30\}$ iterations. For faster convergence, the update (16) is repeated for 10 times at each AO iteration.

The first numerical example is shown in Fig. 1, where the MSE error is plotted against the dimension of the state space N, under 1000 Monte Carlo simulations. We first observe that the offset-unaware's MSE can be $10^4 - 10^5$ times of the offset-free's one. This shows the detrimental effect of sampling offset to state estimation. Secondly, applying the proposed algorithm can significantly improve the MSE in the presence of asynchronous sampling. It is important to point out that from Corollary 1, the state and timing errors are uniquely identified only if $KM \geq 2N$. In this example, we require that $N \leq 6$. This explains the increasing slope in the MSE when $N > 6$.

The next scenario of interest is the state estimation problem in WAMS. Here, we consider the IEEE-14 bus system in MATPOWER [12] as illustrated by Fig. 2. The PMUs are placed on 7 buses in the system and the corresponding dimensions are $\sum_{i=1}^{K} M_i = 31$ and $N = 14$. To obtain the spectrum $\mathbf{X}(e^{j\omega})$, We generate L samples of the complex voltage $\mathbf{x}(t) = \mathbf{v}(t)$, taken at an interval of 10 minutes. The offset error b_i is uniformly distributed over $[-0.5, 0.5]$. The decentralized version for these algorithms based on Gossip exchange in Section III-A is also illustrated. For comparison purpose, we have implemented the diffusion LMS [10] for tackling (15). Here, we assume that the underlying communication networks between PMUs is fully connected. For the decentralized algorithms, the MSE is calculated as $\frac{1}{KL}\sum_{i=1}^{K}\sum_{\omega\in\Omega}\|\hat{\mathbf{X}}_i^{(k)}(e^{j\omega}) - \mathbf{X}(e^{j\omega})\|_2^2$.

The second numerical example demonstrates the MSE performance of the joint state and timing regression problem against frame size L under 100 Monte Carlo simulations; as shown in Fig. 3. Firstly, let us focus on the proposed algorithm (centralized). Especially, the performance is better than the results obtained with synthetic data in Fig. 1. A possible explanation for this is that the measurement matrix $\mathbf{H}_i$ in WAMS is sparse. Secondly, the performance with the gossip-based implementation is also shown. As observed, the proposed gossip-based algorithm achieves a similar performance as its centralized counterpart. For a smaller number of iterations, its performance is also significantly better than its LMS counterpart. However, we remark that in the LMS implementation of (15), the sensors are not required to compute any matrix inverses. The LMS algorithm is thus more suitable for large-scale systems.

To summarize, in this paper, we have considered the state estimation problem with sampling offsets. We have developed a new formulation for the problem and provided a necessary and sufficient condition for the recovery of *both* timing error and state to be done. Moreover, we proposed an algorithm that tackles the joint regression problem and showed that the algorithm can be implemented in a decentralized manner. Furthermore, the effectiveness of the proposed algorithm is corroborated by our simulation results.

References

[1] X. Li and A. Scaglione, "Robust Decentralized State Estimation and Tracking for Power Systems via Network Gossiping," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1184–1194, Jul. 2013.

[2] L. Xie, D.-h. Choi, S. Kar, and H. V. Poor, "Fully Distributed State Estimation for Wide-Area Monitoring Systems," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1154–1169, Sep. 2012.

[3] D. P. Shepard, T. E. Humphreys, and A. Fansler, "Evaluation of the vulnerability of phasor measurement units to GPS spoofing attacks," *International Journal of Critical Infrastructure Protection*, vol. 5, no. 3-4, pp. 146–153, Dec. 2012.

[4] J.-O. Nilsson, I. Skog, and P. Handel, "Joint state and measurement time-delay estimation of nonlinear state space systems," in *Proc' ISSPA 2010*, May 2010, pp. 324–328.

[5] S. Julier and J. Uhlmann, "Fusion of time delayed measurements with uncertain time delays," in *Proc' ACC 2005*, pp. 4028–4033.

[6] S. Huang and B. C. Levy, "Blind Calibration of Timing Offsets for Four-Channel Time-Interleaved ADCs," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 54, no. 4, pp. 863–876, Apr. 2007.

[7] J. Goodman, B. Miller, M. Herman, G. Raz, and J. Jackson, "Polyphase Nonlinear Equalization of Time-Interleaved Analog-to-Digital Converters," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 3, pp. 362–373, Jun. 2009.

[8] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, Sep. 1999.

[9] S. Bonettini, "Inexact block coordinate descent methods with application to non-negative matrix factorization," *IMA Journal of Numerical Analysis*, vol. 31, no. 4, pp. 1431–1452, 2011.

[10] S.-Y. Tu and A. H. Sayed, "Diffusion Strategies Outperform Consensus Strategies for Distributed Estimation Over Adaptive Networks," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6217–6234, Dec. 2012.

[11] A. Nedic, A. Ozdaglar, and P. Parrilo, "Constrained Consensus and Optimization in Multi-Agent Networks," *IEEE Trans. Autom. Control*, vol. 55, no. 4, pp. 922–938, Apr. 2010.

[12] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.