On Clock Synchronization Algorithms for Wireless Sensor Networks Under Unknown Delay

Mei Leng and Yik-Chung Wu

Abstract—In this paper, three clock-synchronization algorithms for wireless sensor networks (WSNs) under unknown delay are derived. They include the maximum-likelihood estimator (MLE), a generalization of the estimator of Noh *et al.*, and a novel low-complexity estimator. Their corresponding performance bounds are derived and compared, and complexities are also analyzed. It is found that the MLE achieves the best performance with the price of high complexity. For the generalized version of the estimator of Noh *et al.*, although it has low complexity, its performance is degraded with respect to the MLE. On the other hand, the newly proposed estimator achieves the same performance as the MLE, and the complexity is at the same level as that of the generalized version of the estimator of Noh *et al.*

Index Terms—Clock synchronization, two-way message exchange mechanism, wireless sensor network (WSN).

I. Introduction

WIRELESS sensor networks (WSNs) have emerged as an important research area in recent years. With their feasibility rapidly growing, WSNs have been regarded as fundamental infrastructures for future ubiquitous communications due to a variety of promising potential applications, such as monitoring of the health status of humans and the environment and control and instrumentation of industrial machines and home appliances [1]–[3]. Most of these applications require the operation of data fusion, power management, and transmission scheduling among a large set of sensor nodes, which, in turn, require all the nodes running on a common time frame. However, every individual sensor in the WSN has its own clock. Different clocks drift from each other over time due to many factors, such as imperfection of the oscillators and environmental changes. This makes clock synchronization between different nodes indispensable.

Clock synchronization is not an easy task in practice due to several unique properties of WSNs [4]. The first and most important property is limited power supply in low-end sensor nodes. In some cases, sensor nodes are not even rechargeable. To save power, synchronization procedures should be simple, and the frequency of resynchronization should be low. This makes simplicity and accuracy the primary concerns of clock synchronization algorithms for WSNs.

The second challenge for clock synchronization is the unknown message delays in the physical and medium-access-control layers. Kopetz and Ochsenreiter [5], for the first time, analyzed the process of message delay and decomposed the unknown delay into different components. Furthermore, these components can be grouped into two portions [7]: 1) the fixed delay (including transmission time, propagation time, and reception time) and 2) the random delay (including send time, access time, and receive time). For the fixed-delay part, if it is not explicitly modeled, it will be treated as part of the time offset, thus lowering the accuracy of timing parameter estimation.

To effectively deal with the clock synchronization of WSNs, many synchronization protocols have been designed over the past few years [6]–[12]. In particular, the timing-sync protocol for sensor networks (TPSN) [7] is one of the most widely used protocols. TPSN is based on the classical two-way message-exchange mechanism and corrects the clock offset between two nodes. Unfortunately, the clock skew is not estimated in TPSN, resulting in frequent resynchronization. Noh *et al.* [15] generalized TPSN to jointly estimate the clock offset and skew, based on an assumption that the fixed delay is known. However, the fixed delay is usually unknown in practice. Although another maximum-likelihood-like estimator (MLLE) that does not require knowledge of the fixed delay was also derived in [15], the performance of this estimator is not satisfactory.

In this paper, based on the two-way message-exchange mechanism, we propose three estimators for the joint estimation of the clock offset and skew without knowledge of the fixed delay. The derived estimator includes the maximum-likelihood estimator (MLE), a generalized estimator of the MLLE, and a newly proposed low-complexity estimator. The rest of this paper is organized as follows: The system model is first introduced in Section II. Three estimators are proposed in Section III, and their performance bounds are derived in Section IV. To compare these estimators, their performance bounds and complexities are compared in Sections V and VI, respectively. Simulation results are presented in Section VII to corroborate the analyses, and finally, conclusions are drawn in Section VIII.

II. System Model

We consider the synchronization between a parent node P and its child node S based on a two-way timing message-exchange mechanism, as shown in Fig. 1. In the ith round of

The authors are with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong (e-mail: meileng@eee.hku.hk; ycwu@eee.hku.hk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.
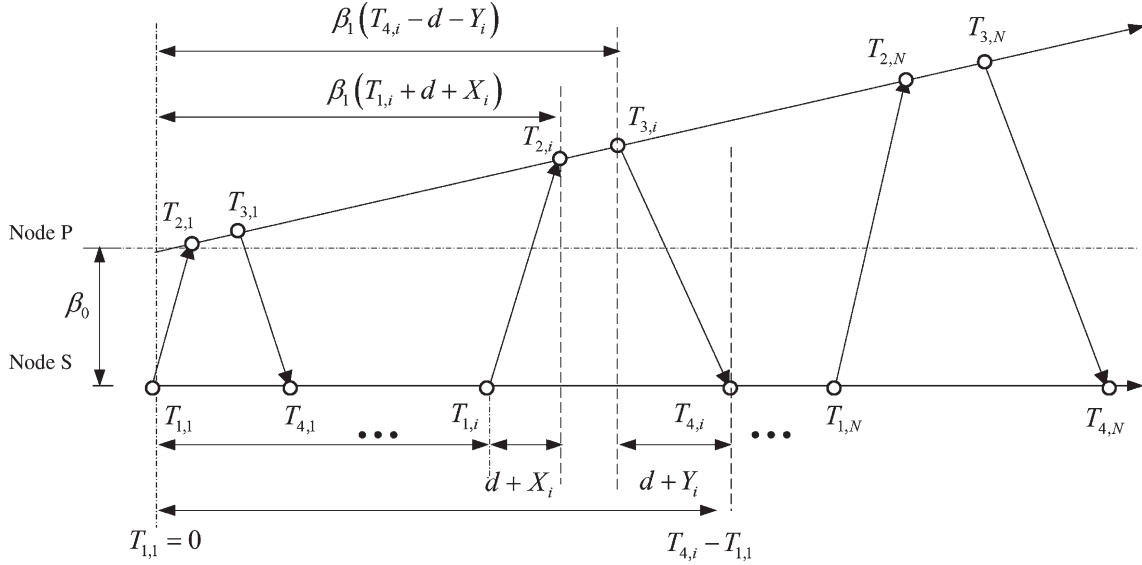
Fig. 1. Two-way time stamps exchange between two nodes S and P.

message exchange, node S sends a synchronization message to node P at $T_{1,i}$. Node P records its time $T_{2,i}$ at the reception of that message and replies to node S at $T_{3,i}$. The replied message contains both time stamps $T_{2,i}$ and $T_{3,i}$. Then, node S records the reception time of node P's reply as $T_{4,i}$. Note that $T_{1,i}$ and $T_{4,i}$ are the time stamps recorded by the clock of node S, whereas $T_{2,i}$ and $T_{3,i}$ are recorded by that of node P. After N rounds of message exchange, node S obtains a set of time stamps $\{T_{1,i}, T_{2,i}, T_{3,i}, T_{4,i}\}_{i=1}^{N}$. The aforementioned procedure can be modeled as [15]

$$T_{2,i} = \beta_1 \times T_{1,i} + \beta_0 + \beta_1 \times (d + X_i) \qquad (1)$$

$$T_{3,i} = \beta_1 \times T_{4,i} + \beta_0 - \beta_1 \times (d + Y_i) \qquad (2)$$

where β_0 and β_1 represent the clock offset and clock skew of node S with respect to node P, respectively; d stands for the fixed portion of message delay from one node to another; and X_i and Y_i are the variable portions of the message delay. In the literature, this random delay has been modeled as random variables following different distributions based on different justifications and applications, such as Gaussian distribution, exponential distribution, and Gamma and Weibull distribution. By considering the random delay as a variable due to numerous independent random processes, we can assume that X_i and Y_i are independent identically distributed (i.i.d.) Gaussian random variables, and this assumption was experimentally verified in [8]. The goal is to estimate clock offset β_0 and clock skew β_1 based on the observation of a set of time stamps $\{T_{1,i}, T_{2,i}, T_{3,i}, T_{4,i}\}_{i=1}^{N}$.

If d is known, it is easy to rewrite the system model (1) and (2) in the standard linear form. Then, β_1 and β_0 can easily be estimated [14]. Noh *et al.* [15] made such assumption and derived the MLE for the joint estimation of clock offset β_0 and clock skew β_1. Unfortunately, in most cases, the value of d will be unknown, as will a parameter that needs to be estimated, such as in the context of node localization. Therefore, efficient methods for estimating β_1 and β_0 with unknown delay d are of great interest. In the following, we will present and analyze

three different methods for estimating β_0 and β_1 when d is unknown.

Remark 1: In the system model, d is assumed to be unknown but of the same value for both the uplink (from node S to node P) and the downlink (from node P to node S). This is because the packets for timing message exchanges between two nodes are usually of the same length and data rate, which makes both nodes undergo similar transmission and reception times. In addition, the propagation time is determined by the distance between two nodes. Since it is assumed that the relative position of two nodes does not change during the message exchange, the propagation time stays the same.

III. ESTIMATION ALGORITHMS FOR THE CLOCK PARAMETERS

A. MLE

Based on the system model (1) and (2), we will derive the MLE for the joint estimation of clock offset β_0, clock skew β_1, and fixed delay d. Rewrite (1) and (2) as

$$\frac{1}{\beta_1} \cdot T_{2,i} = T_{1,i} + \frac{\beta_0}{\beta_1} + d + X_i \qquad (3)$$

$$\frac{1}{\beta_1} \cdot T_{3,i} = T_{4,i} + \frac{\beta_0}{\beta_1} - d - Y_i. \qquad (4)$$

Defining $\theta_1 = 1/\beta_1$ and $\theta_0 = \beta_0/\beta_1$ and stacking all the time stamps in matrix form, we have

$$\underbrace{\begin{bmatrix} T_{1,1} \\ \vdots \\ T_{1,N} \\ -T_{4,1} \\ \vdots \\ -T_{4,N} \end{bmatrix}}_{\triangleq \mathbf{T}_S} + d \cdot \mathbf{1} = \underbrace{\begin{bmatrix} T_{2,1} & -1 \\ \vdots & \vdots \\ T_{2,N} & -1 \\ -T_{3,1} & 1 \\ \vdots & \vdots \\ -T_{3,N} & 1 \end{bmatrix}}_{\triangleq \mathbf{T}_P} \underbrace{\begin{bmatrix} \theta_1 \\ \theta_0 \end{bmatrix}}_{\triangleq \Theta} - \underbrace{\begin{bmatrix} X_1 \\ \vdots \\ X_N \\ Y_1 \\ \vdots \\ Y_N \end{bmatrix}}_{\triangleq \mathbf{Z}} \qquad (5)$$

where $\mathbf{1} = [1, \ldots, 1]^T$ with dimension $2N \times 1$.

Since the random delays X_i and Y_i follow i.i.d. Gaussian distribution [i.e., $X_i \sim \mathcal{N}(0, \sigma^2)$, $Y_i \sim \mathcal{N}(0, \sigma^2)$], the probability density function (pdf) of the observed time stamps conditioned on d and $\boldsymbol{\Theta}$ is

$$\ln f(\mathbf{T}_S, \mathbf{T}_P; \boldsymbol{\Theta}, d) = \ln \frac{N}{2\pi\sigma^2} - \frac{\|\mathbf{T}_S + d\mathbf{1} - \mathbf{T}_P\boldsymbol{\Theta}\|^2}{2\sigma^2}. \quad (6)$$

For a fixed d, the MLE of $\boldsymbol{\Theta}$ is [14]

$$\hat{\boldsymbol{\Theta}}(d) = \left(\mathbf{T}_P^H \mathbf{T}_P\right)^{-1} \mathbf{T}_P^H (\mathbf{T}_S + d\mathbf{1}). \quad (7)$$

Plugging (7) into (6) and ignoring the irrelevant constants, we arrive at the compressed likelihood function with only one parameter d. Denoting $\mathbf{P} = \mathbf{I}_{2N} - \mathbf{T}_P(\mathbf{T}_P^H \mathbf{T}_P)^{-1}\mathbf{T}_P^H$ and noting that $\mathbf{P}^H\mathbf{P} = \mathbf{P}$, the likelihood function is given by

$$\Lambda(d) = \|\mathbf{P}(\mathbf{T}_S + d\mathbf{1})\|^2. \quad (8)$$

Taking the derivative over the likelihood function (8) with respect to d and setting the result to zero, it can easily be shown that the estimator for d is

$$\hat{d} = -\frac{1}{2} \cdot \frac{\mathbf{1}^H \mathbf{P} \mathbf{T}_S + \mathbf{T}_S^H \mathbf{P}\mathbf{1}}{\mathbf{1}^H \mathbf{P}\mathbf{1}}. \quad (9)$$

Putting the estimated $\hat{d}$ back into (7), we finally get the estimator of β_1 and β_0 as $\hat{\beta}_1 = 1/[\hat{\boldsymbol{\Theta}}(\hat{d})]_1$ and $\hat{\beta}_0 = [\hat{\boldsymbol{\Theta}}(\hat{d})]_2/[\hat{\boldsymbol{\Theta}}(\hat{d})]_1$, respectively, where $[\mathbf{v}]_i$ denotes the ith element of vector $\mathbf{v}$.

B. Generalization of the MLLE of [15]

By observing that the clock difference between two nodes is monotonically increasing and every message exchange experiences the same amount of clock offset β_0 and fixed delay d, Noh *et al.* [15] proposed subtracting $T_{r,1}$ from $T_{r,N}$ to obtain $W_r = T_{r,N} - T_{r,1}(r = \{1, 2, 3, 4\})$, such that W_r does not depend on β_0 and d. Then, estimation of β_1 based on $\{W_1, W_2, W_3, W_4\}$ can easily be derived. This estimator was named MLLE. One drawback of this estimator is that only the time stamps in the first and last rounds of message exchange are used, and not all observations have been exploited. The MLLE will be generalized here by using all the available time stamps.

First, define $D_{r,j} \triangleq T_{r,\alpha+j} - T_{r,j}(j = 1, \dots, N - \alpha$ and $r = 1, 2, 3, 4)$, where $\alpha \in \{1, \dots, N - 1\}$ indicates the gap between two subtracting time stamps. Apparently, to construct the sequence $\{D_{r,j}\}$, the parent node P has to perform at least two rounds of message exchange with its child node S, i.e., $N \geq 2$. Using (1) and (2), it can easily be shown that the following relationships hold:

$$D_{2,j} = \beta_1 D_{1,j} + \beta_1(X_{\alpha+j} - X_j) \quad (10)$$

$$D_{3,j} = \beta_1 D_{4,j} - \beta_1(Y_{\alpha+j} - Y_j) \quad (11)$$

for $j = 1, \dots, N - \alpha$. Dividing (10) and (11) by β_1 and stacking all the $D_{r,j}$ in matrix form, we have

$$\begin{bmatrix} -D_{1,1} \\ \vdots \\ -D_{1,N-\alpha} \\ D_{4,1} \\ \vdots \\ D_{4,N-\alpha} \end{bmatrix} = \begin{bmatrix} -D_{2,1} \\ \vdots \\ -D_{2,N-\alpha} \\ D_{3,1} \\ \vdots \\ D_{3,N-\alpha} \end{bmatrix} \cdot \theta_1 + \begin{bmatrix} X_{\alpha+1} - X_1 \\ \vdots \\ X_N - X_{N-\alpha} \\ Y_{\alpha+1} - Y_1 \\ \vdots \\ Y_N - Y_{N-\alpha} \end{bmatrix}. \quad (12)$$

Since X_i and Y_i are i.i.d. Gaussian distributed with variance σ^2, $(X_{\alpha+i} - X_i) \sim \mathcal{N}(0, 2\sigma^2)$, and $(Y_{\alpha+i} - Y_i) \sim \mathcal{N}(0, 2\sigma^2)$. From (12), the MLE for β_1 is given by [14]

$$\hat{\beta}_1 = \frac{1}{\hat{\theta}_1} = \frac{\sum_{j=1}^{N-\alpha} \left(D_{2,j}^2 + D_{3,j}^2\right)}{\sum_{j=1}^{N-\alpha} \left(D_{1,j}D_{2,j} + D_{4,j}D_{3,j}\right)}. \quad (13)$$

To estimate the clock offset β_0, we plug the estimated clock skew back into the original model (1) and (2) and obtain

$$T_{2,i} - \hat{\beta}_1 \times T_{1,i} = \beta_0 + \hat{\beta}_1 \times (d + X_i) \quad (14)$$

$$T_{3,i} - \hat{\beta}_1 \times T_{4,i} = \beta_0 - \hat{\beta}_1 \times (d + Y_i). \quad (15)$$

Adding (14) and (15) together, we have $(T_{2,i} + T_{3,i}) - \hat{\beta}_1(T_{1,i} + T_{4,i}) = 2\beta_0 + \hat{\beta}_1(X_i - Y_i)$. Since $(X_i - Y_i) \sim \mathcal{N}(0, 2\sigma^2)$, the clock offset can be estimated as

$$\hat{\beta}_0 = \frac{1}{2N} \sum_{i=1}^{N} \left[(T_{2,i} + T_{3,i}) - \hat{\beta}_1(T_{1,i} + T_{4,i})\right]. \quad (16)$$

Remark 2: When $\alpha = N - 1$, j must be 1, and we have $D_{r,1} = T_{r,N} - T_{r,1} = W_r$. Thus, the estimator (13) reduces to

$$\hat{\beta}_1 = \frac{W_2^2 + W_3^2}{W_1 W_2 + W_3 W_4} \quad (17)$$

which is exactly the estimator proposed by Noh *et al.* [15].

C. Proposed Low-Complexity Estimator

By observing that the uplink and downlink undergo the same amount of fixed delay, we can rewrite the original model by adding (1) to (2), and the modified model is given by

$$T_{2,i} + T_{3,i} = \beta_1 \times (T_{1,i} + T_{4,i}) + 2\beta_0 + \beta_1 \times (X_i - Y_i). \quad (18)$$

Dividing the aforementioned equation by β_1 and stacking all the time stamps in matrix form, the model becomes

$$\underbrace{\begin{bmatrix} T_{1,1} + T_{4,1} \\ \vdots \\ T_{1,N} + T_{4,N} \end{bmatrix}}_{\triangleq \mathbf{T}_{S'}} = \underbrace{\begin{bmatrix} T_{2,1} + T_{3,1} & -2 \\ \vdots & \vdots \\ T_{2,N} + T_{3,N} & -2 \end{bmatrix}}_{\triangleq \mathbf{T}_{P'}} \underbrace{\begin{bmatrix} \theta_1 \\ \theta_0 \end{bmatrix}}_{=\boldsymbol{\Theta}} + \underbrace{\begin{bmatrix} Y_1 - X_1 \\ \vdots \\ Y_N - X_N \end{bmatrix}}_{\triangleq \mathbf{Z'}}. \quad (19)$$

Since the i.i.d. random delay X_i and Y_i follow Gaussian distribution with variance σ^2, $Z_i' \triangleq (Y_i - X_i) \sim \mathcal{N}(0, 2\sigma^2)$. The MLE for $\boldsymbol{\Theta}$ is given by

$$\hat{\boldsymbol{\Theta}} = \left(\mathbf{T}_{P'}^H \mathbf{T}_{P'}\right)^{-1} \mathbf{T}_{P'}^H \mathbf{T}_{S'}. \tag{20}$$

The estimators for β_1 and β_0 are given by $\hat{\beta}_1 = 1/[\hat{\boldsymbol{\Theta}}]_1$ and $\hat{\beta}_0 = [\hat{\boldsymbol{\Theta}}]_2/[\hat{\boldsymbol{\Theta}}]_1$, respectively.

IV. PERFORMANCE BOUND DERIVATION

A. Cramer-Rao Lower Bound (CRLB)

Noh *et al.* [15] presented the CRLB for the joint estimation of clock skew β_1 and clock offset β_0 under the assumption that the fixed delay d is known. When the fixed delay is included as a parameter, we need to rederive the CRLB for each parameter in the joint estimation.

The PDF in (6) can be rewritten as

$$\ln f\left(\{T_{1,i}, T_{2,i}, T_{3,i}, T_{4,i}\}_{i=1}^N; \beta_1, \beta_0, d\right) = \ln \frac{N}{2\pi\sigma^2} - \frac{1}{2\sigma^2}$$
$$\times \sum_{i=1}^N \left[\left(\frac{T_{2,i}}{\beta_1} - T_{1,i} - \frac{\beta_0}{\beta_1} - d\right)^2 + \left(T_{4,i} - \frac{T_{3,i}}{\beta_1} + \frac{\beta_0}{\beta_1} - d\right)^2\right]. \tag{21}$$

Then, the Fisher information matrix is given by [14]

$$\text{FIM}(\beta_1, \beta_0, d) = \begin{bmatrix} -\mathbf{E}\frac{\partial^2 \ln f}{\partial \beta_1^2} & -\mathbf{E}\frac{\partial^2 \ln f}{\partial \beta_1 \partial \beta_0} & -\mathbf{E}\frac{\partial^2 \ln f}{\partial \beta_1 \partial d} \\ -\mathbf{E}\frac{\partial^2 \ln f}{\partial \beta_0^2} & -\mathbf{E}\frac{\partial^2 \ln f}{\partial \beta_0 \partial \beta_1} & -\mathbf{E}\frac{\partial^2 \ln f}{\partial \beta_0 \partial d} \\ -\mathbf{E}\frac{\partial^2 \ln f}{\partial d^2} & -\mathbf{E}\frac{\partial^2 \ln f}{\partial d \partial \beta_1} & -\mathbf{E}\frac{\partial^2 \ln f}{\partial d \partial \beta_0} \end{bmatrix}$$
$$= \begin{bmatrix} \mathcal{A} & \mathcal{B} & \mathcal{C} \\ \mathcal{B} & \frac{2N}{\beta_1^2} & 0 \\ \mathcal{C} & 0 & 2N \end{bmatrix} \tag{22}$$

where

$$\mathcal{A} \triangleq \frac{1}{\beta_1^4} \sum_{i=1}^N \left[\beta_1^2(T_{1,i} + d)^2 + \beta_1^2\sigma^2 + (T_{3,i} - \beta_0)^2\right] \tag{23}$$

$$\mathcal{B} \triangleq \frac{1}{\beta_1^3} \sum_{i=1}^N \left[\beta_1(T_{1,i} + d) + (T_{3,i} - \beta_0)\right] \tag{24}$$

$$\mathcal{C} \triangleq \frac{1}{\beta_1^2} \sum_{i=1}^N \left[\beta_1(T_{1,i} + d) - (T_{3,i} - \beta_0)\right]. \tag{25}$$

By inverting the matrix (22), it can be shown that the CRLB for each parameter is

$$\text{CRLB}(\beta_1) = \frac{2N\sigma^2}{2N\mathcal{A} - \beta_1^2\mathcal{B}^2 - \mathcal{C}^2} \tag{26}$$

$$\text{CRLB}(\beta_0) = \frac{\sigma^2\beta_1^2(2N\mathcal{A} - \mathcal{C}^2)}{2N\left(2N\mathcal{A} - \beta_1^2\mathcal{B}^2 - \mathcal{C}^2\right)} \tag{27}$$

$$\text{CRLB}(d) = \frac{\sigma^2\left(2N\mathcal{A} - \beta_1^2\mathcal{B}^2\right)}{2N\left(2N\mathcal{A} - \beta_1^2\mathcal{B}^2 - \mathcal{C}^2\right)}. \tag{28}$$

Remark 3: In the two-way message exchange mechanism, the values of time stamps $T_{1,i}$ and $T_{3,i}$ are determined by the two nodes and should be seen as the system inputs, with $T_{2,i}$ and $T_{4,i}$ being the corresponding outputs. Once the inputs and statistics of the random delay are fixed, the CRLB should be fixed and independent of the system outputs. Therefore, the CRLB previously derived depends only on time stamps $\{T_{1,i}, T_{3,i}\}_{i=1}^N$ and the random-delay variance. However, the CRLB derived in [15] was based on all the time stamps $\{T_{1,i}, T_{2,i}, T_{3,i}, T_{4,i}\}_{i=1}^N$, and CRLB may change according to different realizations of $\{T_{2,i}, T_{4,i}\}_{i=1}^N$.

B. Performance Bound for the Generalized MLLE

Strictly speaking, for a set of data generated from a specific system model, there should be a unique and lowest performance bound for any estimator dealing with these data. The CRLB given in Section IV-A is just such a lowest bound for unbiased estimators. However, if the data are manipulated in the process of estimation such that the system model for estimation is changed, we need to know whether the manipulation causes degradation of the performance limit. Therefore, a performance bound for the generalized MLLE is needed.

Since (12) is a standard linear model and the random delays are i.i.d. Gaussian distributed, following the standard method of deriving CRLB [14], the performance bound for β_1 can be shown to be

$$\text{PB}_g(\beta_1) = \frac{2\sigma^2\beta_1^4}{\sum_{j=1}^{N-\alpha}\left(\beta_1^2 D_{1,j}^2 + D_{3,j}^2 + 6\beta_1^2\sigma^2\right)}. \tag{29}$$

Furthermore, from the estimator for β_0 in (16), together with (1) and (2), the estimation error for the clock offset is given by

$$\hat{\beta}_0 - \beta_0 = \frac{\hat{\beta}_1 - \beta_1}{2N}\sum_{i=1}^N(T_{1,i} + T_{4,i}) + \frac{\beta_1}{2N}\sum_{i=1}^N(X_i - Y_i). \tag{30}$$

Computing the variance on both sides of the aforementioned equation, it can be shown that

$$\text{PB}_g(\beta_0) = \frac{\sigma^2\beta_1^2}{2N} + \frac{\text{PB}_g(\beta_1)}{4N^2}$$
$$\times \left\{\frac{1}{\beta_1^2}\left[\sum_{i=1}^N \beta_1(T_{1,i} + d) + (T_{3,i} - \beta_0)\right]^2 + N\sigma^2\right\} \tag{31}$$

where we use $\mathbf{var}X = \mathbf{E}X^2 - [\mathbf{E}X]^2$ and the fact that the estimator for the clock skew β_1 is asymptotically unbiased.

Clearly, from (29) and (31), the performance bounds of the generalized MLLE for both β_1 and β_0 depend on the integer parameter α. To obtain the best performance, this user-defined α has to be carefully chosen. As shown in Appendix A, the optimal α in the sense of minimizing $\text{PB}_g(\beta_1)$ and $\text{PB}_g(\beta_0)$ is given by the following proposition:

Proposition 1 (Choosing the Optimal α): Assuming $T_{1,i} = iH$ and $T_{3,i} = iG$, where H and G are the gap between neighboring time stamps in $T_{1,i}$ and $T_{3,i}$, respectively, the optimal α

in the sense of minimizing both $\mathrm{PB}_g(\beta_0)$ and $\mathrm{PB}_g(\beta_1)$, which is denoted as α^*, is given by

$$\alpha^* = \arg \max_{\alpha = \lfloor \alpha_r^* \rfloor, \lceil \alpha_r^* \rceil} \Phi(\alpha) \qquad (32)$$

where $\Phi(\alpha) = (\beta_1^2 H^2 + G^2)(-\alpha^3 + N\alpha^2) + 6(N-\alpha)\beta_1^2\sigma^2$ and

$$\alpha_r^* = \begin{cases} \frac{1}{3}N + \sqrt{\frac{1}{9}N^2 - \frac{2\beta_1^2\sigma^2}{\beta_1^2 H^2 + G^2}}, & \sigma^2 \le \lambda_{Th} \text{ or } N = 3 \\ 1, & \sigma^2 > \lambda_{Th} \text{ and } N \ne 3 \end{cases} \qquad (33)$$

where $\lambda_{Th} \triangleq (2/27N^3 - N + 1)(\beta_1^2 H^2 + G^2)/[2\beta_1^2(N - 3)]$. Moreover, in the high-SNR region, the optimal α reduces to $\alpha^* = 2k + \lceil j/2 \rceil$, where k and j are related to N by $N = 3k + j$, with $k \in \{0, 1, 2, 3, \ldots\}$ and $j \in \{0, 1, 2\}$.

Remark 3: Since the values of $T_{1,i}$ and $T_{3,i}$ are determined by the two nodes in the two-way message exchange mechanism, the assumption $T_{1,i} = iH$ and $T_{3,i} = iG$ is practical and causes no loss of generality. Furthermore, as numerically shown in Section VII, $\alpha^* = 2k + \lceil j/2 \rceil$ provides the optimal performance, even when the transmitted time stamps are not uniformly spaced.

C. Performance Bound for the Proposed Low-Complexity Estimator

For the same reason stated in Section IV-B, we need to derive the performance bound for the novel low-complexity estimator. From the model $\mathbf{T}_{S'} = \mathbf{T}_{P'}\boldsymbol{\Theta} + \mathbf{Z}'$ in (19), the pdf of $\mathbf{T}_{S'}$ is $\ln f(\mathbf{T}_{S'}; \boldsymbol{\Theta}) = -N/2 \ln(2\pi\sigma^2) - 1/2\sigma^2 \|\mathbf{T}_{S'} - \mathbf{T}_{P'}\boldsymbol{\Theta}\|^2$. Then, the Fisher information matrix can be derived as [14]

$$\mathrm{FIM}(\beta_1, \beta_0) = \frac{1}{2\sigma^2} \begin{bmatrix} \mathcal{K} & 2\mathcal{B} \\ 2\mathcal{B} & \frac{4N}{\beta_1^2} \end{bmatrix} \qquad (34)$$

where $\mathcal{B}$ was defined in (24), and

$$\mathcal{K} = \frac{1}{\beta_1^2} \sum_{i=1}^{N} \left\{ \left[(T_{1,i} + d) + \frac{1}{\beta_1}(T_{3,i} - \beta_0) \right]^2 + 3\sigma^2 \right\}. \qquad (35)$$

Inverting the 2×2 matrix in (34), the performance bounds for the clock skew β_1 and the clock offset β_0 are given by

$$\mathrm{PB}_P(\beta_1) = \frac{2N\sigma^2}{N\mathcal{K} - \beta_1^2\mathcal{B}^2} \qquad (36)$$

$$\mathrm{PB}_P(\beta_0) = \frac{\sigma^2\beta_1^2\mathcal{K}}{2N\mathcal{K} - 2\beta_1^2\mathcal{B}^2}. \qquad (37)$$

V. FURTHER PERFORMANCE ANALYSES

To see whether the performance limits of the two estimators in Sections III-B and C can meet the CRLB, in this section, we compare their performance bounds with the CRLB. To get insightful conclusions, it is assumed that $T_{1,i} = iH$ and $T_{3,i} = iG$, where H and G are the gap between neighboring time stamps in $T_{1,i}$ and $T_{3,i}$, respectively. Simulation results will be presented in Section VII to corroborate the results in this section when $T_{1,i}$ and $T_{3,i}$ are not uniformly spaced.

A. CRLB and Performance Bound of the Proposed Estimator PB_p

From (26) and (36), the normalized difference between the performance bound $\mathrm{PB}_p(\beta_1)$ of the proposed low-complexity estimator and $\mathrm{CRLB}(\beta_1)$ can be computed as

$$\frac{\mathrm{PB}_P(\beta_1) - \mathrm{CRLB}(\beta_1)}{\mathrm{CRLB}(\beta_1)} = \frac{2N\mathcal{A} - \mathcal{C}^2 - N\mathcal{K}}{N\mathcal{K} - \beta_1^2\mathcal{B}^2}. \qquad (38)$$

Using the definition of $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{K}$ in (23)–(25) and (35), and after some straightforward manipulations, it can be shown that

$$2N\mathcal{A} - \mathcal{C}^2 - N\mathcal{K} = \frac{N^2}{\beta_1^2} \left[\frac{1}{N}\sum_{i=1}^{N} V_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N} V_i \right)^2 - \sigma^2 \right] \qquad (39)$$

$$N\mathcal{K} - \beta_1^2\mathcal{B}^2 = \frac{N^2}{\beta_1^2} \left[\frac{1}{N}\sum_{i=1}^{N} U_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N} U_i \right)^2 + 3\sigma^2 \right] \qquad (40)$$

where $V_i = (T_{1,i} + d) - (1/\beta_1)(T_{3,i} - \beta_0)$, and $U_i = (T_{1,i} + d) + (1/\beta_1)(T_{3,i} - \beta_0)$. With the assumption $T_{1,i} = iH$ and $T_{3,i} = iG$, (39) and (40) can further be simplified as

$$2N\mathcal{A} - \mathcal{C}^2 - N\mathcal{K}$$
$$= \frac{N^2}{\beta_1^4} \left\{ (\beta_1 H - G)^2 \left[\frac{1}{N}\sum_{i=1}^{N} i^2 - \frac{1}{N^2}\left(\sum_{i=1}^{N} i \right)^2 \right] - \beta_1^2\sigma^2 \right\} \qquad (41)$$

$$N\mathcal{K} - \beta_1^2\mathcal{B}^2$$
$$= \frac{N^2}{\beta_1^4} \left\{ (\beta_1 H + G)^2 \left[\frac{1}{N}\sum_{i=1}^{N} i^2 - \frac{1}{N^2}\left(\sum_{i=1}^{N} i \right)^2 \right] + 3\beta_1^2\sigma^2 \right\}. \qquad (42)$$

Putting (41) and (42) into (38) and recognizing that $\sum_{i=1}^{N} i^2/N - (\sum_{i=1}^{N} i)^2/N^2 = (N^2 - 1)/12$, the normalized difference between $\mathrm{PB}_p(\beta_1)$ and $\mathrm{CRLB}(\beta_1)$ is given by

$$\frac{\mathrm{PB}_P(\beta_1) - \mathrm{CRLB}(\beta_1)}{\mathrm{CRLB}(\beta_1)} = \frac{(N^2 - 1)(\beta_1 H - G)^2 - 12\beta_1^2\sigma^2}{(N^2 - 1)(\beta_1 H + G)^2 + 36\beta_1^2\sigma^2} \qquad (43)$$

$$< \frac{(\beta_1 H - G)^2}{(\beta_1 H + G)^2}. \qquad (44)$$

Since H and G are usually of similar values to keep the two-way message exchange in order and the clock skew is generally close to 1, $(\beta_1 H - G)$ is a small value, and the normalized difference is close to zero. For example, supposing $\beta_1 = 0.95$, $H = 25$, and $G = 30$, the normalized difference computed using (44) is approximately 0.0135. Thus, the performance limit of the proposed estimator for clock skew β_1 is almost the same as that of the CRLB. Simulation results shown in Section VII further verify that the performance bound $\mathrm{PB}_p(\beta_1)$ of the proposed estimator is indistinguishable from $\mathrm{CRLB}(\beta_1)$.

Similarly, using (27) and (37), the normalized difference between $\text{PB}_p(\beta_0)$ and $\text{CRLB}(\beta_0)$ can be shown to be

$$\frac{\text{PB}_P(\beta_0) - \text{CRLB}(\beta_0)}{\text{CRLB}(\beta_0)} = \frac{\beta_1^2 \mathcal{B}^2 (2N\mathcal{A} - \mathcal{C}^2 - N\mathcal{K})}{(2N\mathcal{A} - \mathcal{C}^2)(N\mathcal{K} - \beta_1^2 \mathcal{B}^2)} \quad (45)$$

$$= \frac{\gamma \cdot \beta_1^2 \mathcal{B}^2}{2N\mathcal{A} - \mathcal{C}^2} \quad (46)$$

where $\gamma \triangleq (2N\mathcal{A} - \mathcal{C}^2 - N\mathcal{K})/(N\mathcal{K} - \beta_1^2 \mathcal{B}^2)$ is the normalized difference between $\text{PB}_p(\beta_0)$ and $\text{CRLB}(\beta_0)$ in (38). With $T_{1,i} = iH$ and $T_{3,i} = iG$, it is shown in Appendix B that $\beta_1^2 \mathcal{B}^2/(2N\mathcal{A} - \mathcal{C}^2)$ can be written in the form of (47), shown at the bottom of the page.

Therefore, putting (47) back into (46), we have

$$\frac{\text{PB}_P(\beta_0) - \text{CRLB}(\beta_0)}{\text{CRLB}(\beta_0)}$$

$$< \gamma \cdot \frac{3(N+1)(\beta_1 H + G)^2}{2(N-1)(\beta_1^2 H^2 + G^2) + 3(N+1)(\beta_1 H + G)^2}$$

$$< \frac{3(N+1)(\beta_1 H - G)^2}{2(N-1)(\beta_1^2 H^2 + G^2) + 3(N+1)(\beta_1 H + G)^2}. \quad (48)$$

With the same reason previously stated, $(\beta_1 H - G)$ is small; therefore, the normalized difference between $\text{PB}_p(\beta_0)$ and $\text{CRLB}(\beta_0)$ is close to zero. For example, supposing $\beta_1 = 0.95$, $H = 25$, $G = 30$, and $N = 6$, the normalized difference computed using (48) is 0.0109. Thus, we can conclude that the proposed estimator for clock offset β_0 has almost the same performance limit as that of MLE, and this is further verified by simulation results.

B. CRLB and Performance Bound of the Generalized MLLE PB_g

Putting $T_{1,i} = iH$ and $T_{3,i} = iG$ into (26) and (29), and after some straightforward manipulations, the normalized difference between the performance bound $\text{PB}_g(\beta_1)$ of the generalized MLLE and $\text{CRLB}(\beta_1)$ is computed as

$$\frac{\text{PB}_g(\beta_1) - \text{CRLB}(\beta_1)}{\text{CRLB}(\beta_1)} = \left[\alpha^2 + \frac{6\beta_1^2 \sigma^2}{\beta_1^2 H^2 + G^2}\right]^{-1}$$

$$\times \left[\frac{N(N^2 - 1)}{6(N - \alpha)} - \alpha^2 - \frac{2N - 3\alpha}{N - \alpha} \frac{2\beta_1^2 \sigma^2}{\beta_1^2 H^2 + G^2}\right]. \quad (49)$$

To gain more insight, we examine the high-SNR scenario (i.e., σ^2 is small), and the normalized difference becomes

$$\frac{\text{PB}_g(\beta_1) - \text{CRLB}(\beta_1)}{\text{CRLB}(\beta_1)} = \frac{N(N^2 - 1)}{6\alpha^2(N - \alpha)} - 1. \quad (50)$$

Moreover, in the high-SNR scenario, the optimal α that minimizes $\text{PB}_g(\beta_1)$ is given by $\alpha^* = 2k + \lceil j/2 \rceil$, where k and

j are related to N by $N = 3k + j$, with $k \in \{0, 1, 2, 3, \ldots\}$ and $j \in \{0, 1, 2\}$. For $N = 2$ and $N = 3$, the optimal α is 1 and 2, respectively. In addition, the normalized difference is zero in these two cases. Otherwise, the normalized difference (assuming that N is a multiple of 3 for simplicity, and therefore, $\alpha^* = 2N/3$) can be shown to be $0.125 - 1.125/N^2$. Therefore, for a large N, $\text{PB}_g(\beta_1)$ will be at least 10% larger than $\text{CRLB}(\beta_1)$. Simulation results in Section VII further verify that the manipulation of time stamps used in the generalized MLLE causes performance loss and degrades the performance bound of β_1 with respect to $\text{CRLB}(\beta_1)$.

Before comparing the performance bound of the generalized MLLE $\text{PB}_g(\beta_0)$ with $\text{CRLB}(\beta_0)$, we first note the following relationships:

$$\text{CRLB}(\beta_0) = \text{CRLB}(\beta_1) \cdot \frac{\beta_1^2 (2N\mathcal{A} - \mathcal{C}^2)}{4N^2} \quad (51)$$

$$\text{PB}_g(\beta_0) = \frac{\sigma^2 \beta_1^2}{2N} + \frac{\text{PB}_g(\beta_1)}{4N^2}\left(\beta_1^4 \mathcal{B}^2 + N\sigma^2\right). \quad (52)$$

Therefore, with $\eta \triangleq (\text{PB}_g(\beta_1) - \text{CRLB}(\beta_1))/\text{CRLB}(\beta_1)$, the normalized difference between $\text{PB}_g(\beta_0)$ and $\text{CRLB}(\beta_0)$ is

$$\frac{\text{PB}_g(\beta_0) - \text{CRLB}(\beta_0)}{\text{CRLB}(\beta_0)} = \frac{\eta \beta_1^4 \mathcal{B}^2 + (\eta + 1)N\sigma^2}{\beta_1^2(2N\mathcal{A} - \mathcal{C}^2)}. \quad (53)$$

In the high-SNR region, σ is small, and we can further simplify (53) as

$$\frac{\text{PB}_g(\beta_0) - \text{CRLB}(\beta_0)}{\text{CRLB}(\beta_0)} = \frac{\eta \cdot \beta_1^2 \mathcal{B}^2}{2N\mathcal{A} - \mathcal{C}^2}. \quad (54)$$

Moreover, with $T_{1,i} = iH$ and $T_{3,i} = iG$, it is shown in Appendix B that

$$\frac{\beta_1^2 \mathcal{B}^2}{2N\mathcal{A} - \mathcal{C}^2} > \frac{3}{5}. \quad (55)$$

Therefore, combining (54) and (55), the normalized difference between $\text{PB}_g(\beta_0)$ and $\text{CRLB}(\beta_0)$ is lower bounded by

$$\frac{\text{PB}_g(\beta_0) - \text{CRLB}(\beta_0)}{\text{CRLB}(\beta_0)} > \frac{3}{5}\eta. \quad (56)$$

That is, for $N = 2$ and $N = 3$, we have $\eta = 0$, and the normalized difference between $\text{PB}_g(\beta_0)$ and $\text{CRLB}(\beta_0)$ is zero from (54). On the other hand, for $N > 3$, by assuming that N is a multiple of 3, for simplicity, we have $\eta = 0.125 - 1.125/N^2$, and the normalized difference (56) is lower bounded by $0.077 - 0.675/N^2$, which indicates that $\text{PB}_g(\beta_0)$ will be at least 7% larger than $\text{CRLB}(\beta_0)$ for a large N. Simulation results in Section VII further verify these analyses.

$$\frac{\beta_1^2 \mathcal{B}^2}{2N\mathcal{A} - \mathcal{C}^2} = \frac{3(N+1)^2(\beta_1 H + G)^2}{2(N^2 - 1)(\beta_1^2 H^2 + G^2) + 3(N+1)^2(\beta_1 H + G)^2 + 24\sigma^2 \beta_1^2} \quad (47)$$

TABLE I
COMPLEXITY COMPARISON

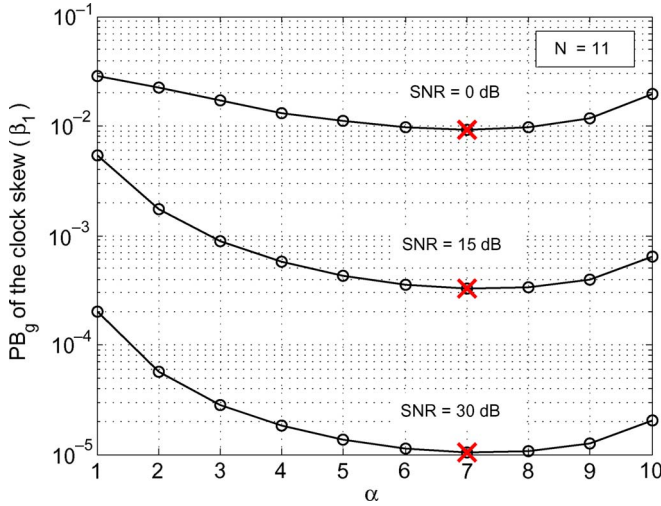| Estimator	parameter	number of +/−	number of ×	number of ÷
MLE	$\beta_0(d)$	$20N - 4$	$16N + 7$	0
	$\beta_1(d)$	$20N - 4$	$16N + 6$	1
	d	$8N^3 + 19N^2 + 14N - 6$	$8N^3 + 11N^2 + 15N + 6$	3
Generalized MLLE	$\beta_0(\beta_1)$	$4N - 1$	2	1
	β_1	$8(N - \alpha) - 2$	$4(N - \alpha)$	1
Proposed Estimator	β_0	$7N - 5$	$5N + 3$	1
	β_1	$7N - 3$	$2N + 4$	1

Fig. 2. Performance bound $PB_g(\beta_1)$ of the estimated clock skew $\hat{\beta}_1$ using the generalized MLLE with different α's. The "x" indicates the value of α calculated by (32), i.e., $\alpha^* = 2k + \lceil j/2 \rceil$, where k and j are related to N by $N = 3k + j$, with $k \in \{0, 1, 2, 3, \ldots\}$ and $j \in \{0, 1, 2\}$.

VI. COMPLEXITY COMPARISONS

To compare the complexity of the three estimators, their number of operations are shown in Table I. The symbol $L(Q)$ has been used to indicate that the calculation of parameter L depends on the precalculated parameter Q. As shown in Table I, for the MLE, β_0 and β_1 have to be calculated based on the value of d, whereas the calculation of d takes complexity order $\mathcal{O}(N^3)$. On the other hand, the other two estimators do not have to calculate d, and the complexity order for the estimation of β_1 and β_0 is only $\mathcal{O}(N)$.

VII. SIMULATION RESULTS AND DISCUSSIONS

Simulations are carried out to compare the performances of the three estimators. The parameters used in the simulation are $d \in (0, 10]$, $\beta_1 \in [0.9, 1.1]$, and $\beta_0 \in [-10, 10]$. In each simulation run, d, β_1, and β_0 are uniformly drawn from their respective ranges. The time stamps at the two nodes are set as $T_{1,i} = iH + \varpi_i$ and $T_{3,i} = iG + \nu_i$, where ϖ_i and ν_i are independent normal random variables with zero mean, and their variances are equal to $0.3H$ and $0.3G$, respectively. The setting for $T_{1,i}$ and $T_{3,i}$ indicates that the message exchange in the simulation follows an unequally spaced transmission. Ten thousand simulation runs were performed to obtain the average performance of each point in the figures.

Fig. 2 shows the performance bound $PB_g(\beta_1)$ of the generalized MLLE for β_1 versus different α's at SNR = 0, 15, and 30 dB, where the SNR is defined as $10 \log[(H^2 +$
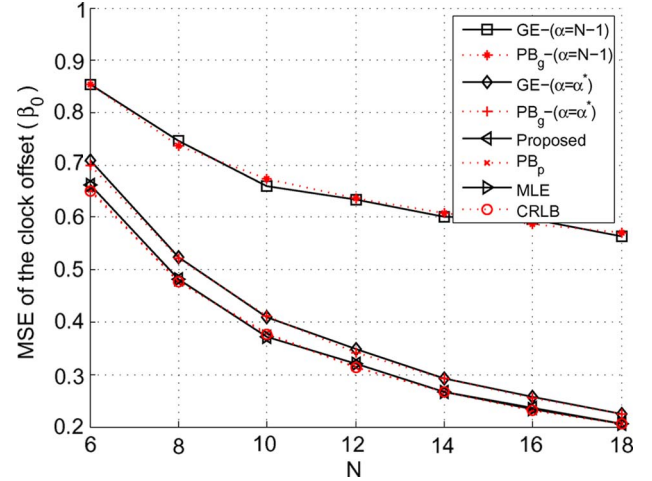

Fig. 3. MSE of estimated clock offset $\hat{\beta}_0$ with respect to the number of time stamps N.

$G^2)/\sigma^2]$. The closed-form optimal α for the high-SNR scenario in (32) is also shown in the figure as an "x." It can be seen from the figure that, even when the variation of the gap between neighboring time stamps is as high as 30%, (32) can still provide the optimal α. Moreover, the value of α calculated by (32) successfully predicts the optimal α not only in the high-SNR scenario but in the low-SNR scenario as well. Since the optimal α for β_1 estimation is also optimal for β_0 estimation, the corresponding figure for β_0 is not shown here.

Fig. 3 shows the MSE for the estimation of the clock offset β_0 using different estimators versus different N's with SNR = 30 dB. In the figure, "GE" stands for the generalized MLLE. In addition, when $\alpha = N - 1$, it corresponds to the original MLLE proposed by Noh *et al.* [15]. As shown in the figure, the performance of the original MLLE significantly deteriorates from the CRLB. When α takes the optimal value, the performance of the generalized MLLE improves but still slightly departs from the CRLB. On the other hand, the proposed estimator achieves the best performance as the MLE, and its performance bound overlaps with CRLB(β_0). Fig. 4 shows the corresponding results for the estimation of the clock skew β_1. The same conclusions as in Fig. 3 can be drawn.

VIII. CONCLUSION

Clock synchronization for WSNs in the presence of unknown Gaussian delay has been discussed based on the two-way message-exchange mechanism. The MLE for the joint estimation of clock skew, clock offset, and fixed delay has first been derived. The MLE achieves good performance, but
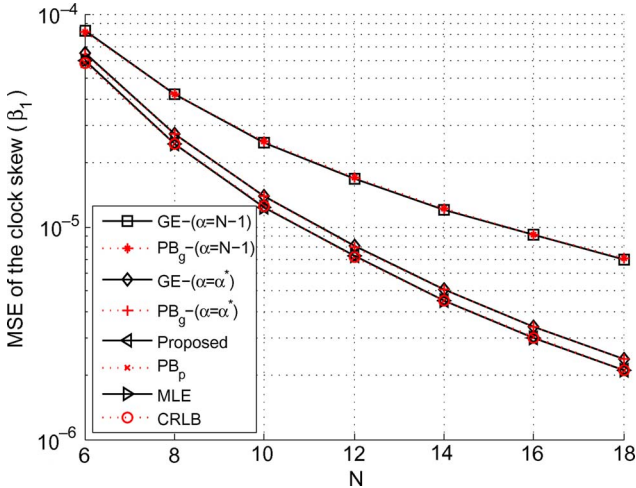
Fig. 4.   MSE of estimated clock skew $\hat{\beta}_1$ with respect to the number of time stamps N.

its complexity is high. Then, two low-complexity estimators have been derived. One is the generalization of the MLLE by Noh *et al.* [15], and the other is a newly proposed estimator. By comparing their performance bounds with the CRLB, it has been found that the proposed low-complexity estimator can achieve CRLB, whereas the generalized MLLE suffers performance degradation. The optimal α for the generalized MLLE has also been derived, and by using the proposed optimal α, the performance degradation can significantly be reduced. Therefore, the generalized MLLE with optimal α proposed in this paper and the newly proposed estimator represent attractive time synchronization algorithms in terms of computational complexities and performances.

APPENDIX A
PROOF OF PROPOSITION 1

With the assumption $T_{1,i} = iH$ and $T_{3,i} = iG$, where H and G are the gap between neighboring time stamps in $T_{1,i}$ and $T_{3,i}$, respectively, and using (1) and (2), the performance bound for β_1 (29) can explicitly be expressed as

$$\mathrm{PB}_g(\beta_1) = \frac{2\sigma^2\beta_1^4}{\left(\beta_1^2 H^2 + G^2\right)\left(-\alpha^3 + N\alpha^2\right) + 6(N-\alpha)\beta_1^2\sigma^2}$$
$$\triangleq \frac{2\sigma^2\beta_1^4}{\Phi(\alpha)}. \tag{57}$$

Finding the optimal α that minimizes $\mathrm{PB}_g(\beta_1)$ is equivalent to solving the following optimization problem:

$$\alpha^* = \arg\max_{\alpha} \Phi(\alpha)$$
$$\text{s.t.} \quad \alpha \in \{1, \dots, N-1\}. \tag{58}$$

First, let us consider a relaxed version of (58), i.e.,

$$\alpha_r^* = \arg\max_{\alpha_r} \Phi(\alpha_r)$$
$$\text{s.t.} \quad 1 \le \alpha_r \le N-1 \tag{59}$$

with α_r being a real number. To solve (59), we look at the convexity/concavity property of the continuous function $\Phi(\alpha_r)$

by computing its second derivative, i.e.,

$$\nabla^2\Phi(\alpha_r) = \left(\beta_1^2 H^2 + G^2\right)(2N - 6\alpha_r).$$

Since β_1, H, and G are all positive, it is easy to find that $\Phi(\alpha_r)$ is convex upward for $\alpha_r \in [1, 1/3N)$ and concave downward for $\alpha_r \in [1/3N, N-1]$.

For $\alpha_r \in [1/3N, N-1]$, the maximum of $\Phi(\alpha_r)$ can be found by setting the first derivative $\nabla\Phi(\alpha_r)$ to zero, and the optimal solution α_r^{*c} in the concave part is given by

$$\alpha_r^{*c} = \frac{1}{3}N + \sqrt{\frac{1}{9}N^2 - \frac{2\beta_1^2\sigma^2}{\beta_1^2 H^2 + G^2}}.$$

For $\alpha_r \in [1, 1/3N]$, the maximum of $\Phi(\alpha_r)$ is achieved at either $\alpha_r = 1$ or $\alpha_r = 1/3N$. When $N = 3$, these two points overlap; therefore, $\alpha_r^* = 1/3N$ in this case. Otherwise, since $\Phi(1) = (N-1)(\beta_1^2 H^2 + G^2 + 6\beta_1^2\sigma^2)$ and $\Phi(1/3N) = 2/27N^3(\beta_1^2 H^2 + G^2) + 4N\beta_1^2\sigma^2$, it can be shown that $\Phi(1)$ is larger than $\Phi(1/3N)$ as long as $\sigma^2 > \lambda_{Th}$, where $\lambda_{Th} \triangleq (2/27N^3 - N + 1)(\beta_1^2 H^2 + G^2)/[2\beta_1^2(N-3)]$. Therefore, the optimal solution α_r^{*v} in the convex part is given by

$$\alpha_r^{*v} = \begin{cases} \frac{1}{3}N, & \sigma^2 \le \lambda_{Th} \text{ or } N = 3 \\ 1, & \sigma^2 > \lambda_{Th} \text{ and } N \ne 3. \end{cases}$$

When $\sigma^2 \le (2/27N^3 - N + 1)(\beta_1^2 H^2 + G^2)/[2\beta_1^2(N-3)]$ is satisfied, it is easy to see that α_r^{*c} always exists and that the inequality $1/3N < \alpha_r^{*c}$ always holds. Since $\Phi(\alpha_r)$ is continuous, we have $\Phi(1/3N) < \Phi(\alpha_r^{*c})$; therefore, considering the whole range for α_r, the optimal solution α_r^* of (59) is

$$\alpha_r^* = \begin{cases} \frac{1}{3}N + \sqrt{\frac{1}{9}N^2 - \frac{2\beta_1^2\sigma^2}{\beta_1^2 H^2 + G^2}}, & \sigma^2 \le \lambda_{Th} \text{ or } N = 3 \\ 1, & \sigma^2 > \lambda_{Th} \text{ and } N \ne 3. \end{cases} \tag{60}$$

Going back to the original problem (58), since the parameter α can only take integer values, the solution of (58) is then given by

$$\alpha^* = \arg\max_{\alpha = \lfloor\alpha_r^*\rfloor, \lceil\alpha_r^*\rceil} \Phi(\alpha). \tag{61}$$

Since $\mathrm{PB}_g(\beta_0)$ depends on α only through $\mathrm{PB}_g(\beta_1)$, as shown in (31), the optimal α for the estimation of β_1 must also be optimal for the estimation of β_0. Therefore, (60) and (61) give the optimal α.

Apparently, α_r^* in (60) depends on σ^2 and β_1, which are difficult or impossible to obtain beforehand. To give a practical solution, we note that, in the high-SNR region (i.e., σ^2 is small), α_r^* is given by $\alpha_r^* = 2/3N$. Expressing $N = 3k + j$, where $k \in \{0, 1, 2, 3, \dots\}$ and $j \in \{0, 1, 2\}$, if $j = 0$, we have $\lfloor\alpha_r^*\rfloor = \lceil\alpha_r^*\rceil = 2k$; thus, $\alpha^* = 2k$. Otherwise, $j \in \{1, 2\}$, and we have $\lfloor\alpha_r^*\rfloor = 2k + j - 1$ and $\lceil\alpha_r^*\rceil = 2k + j$. It can be shown that

$$\Phi(\lfloor\alpha_r^*\rfloor) - \Phi(\lceil\alpha_r^*\rceil) = \left(\beta_1^2 H^2 + G^2\right)\left[(j-1)^2 + k(2j-3)\right] \tag{62}$$

$$\begin{cases} < 0 & j = 1 \\ > 0 & j = 2. \end{cases} \tag{63}$$

Therefore, the optimal value for α is given by

$$\alpha^* = \begin{cases} 2k, & j = 0 \\ 2k + j, & j = 1 \\ 2k + j - 1, & j = 2. \end{cases} \tag{64}$$

In summary, the optimal α can compactly be expressed as $\alpha^* = 2k + \lceil j/2 \rceil$ in the high-SNR region.

APPENDIX B
PROOF OF (47) AND (55)

Under the assumption $T_{1,i} = iH$ and $T_{3,i} = iG$, we have

$$\beta_1^2 \mathcal{B}^2 = \frac{1}{4\beta_1^4}[N(N+1)(\beta_1 H + G) + 2(\beta_1 d - \beta_0)]^2 \tag{65}$$

$$2N\mathcal{A} - \mathcal{C}^2 = \frac{1}{6\beta_1^4}N^2(N^2-1)\left(\beta_1^2 H^2 + G^2\right)$$
$$+ \frac{1}{4\beta_1^4}[N(N+1)(\beta_1 H + G) + 2(\beta_1 d - \beta_0)]^2$$
$$+ \frac{2N^2\sigma^2}{\beta_1^2}. \tag{66}$$

Comparing the two terms $N(N-1)(\beta_1 H + G)$ and $2(\beta_1 d - \beta_0)$, since the first term includes a factor on the order N^2, the second term $2(\beta_1 d - \beta_0)$ is much smaller than $N(N+1)(\beta_1 H + G)$ and can be neglected. Computing the ratio between $\beta_1^2 \mathcal{B}^2$ and $2N\mathcal{A} - \mathcal{C}^2$, we have (47).

Moreover, in the high-SNR scenario, σ is small; thus, we can further neglect the term $24\sigma^2 \beta_1^2$ in (47) and have the following equality:

$$\frac{\beta_1^2 \mathcal{B}^2}{2N\mathcal{A} - \mathcal{C}^2} = \frac{3}{3 + 2(N-1)(\beta_1^2 H^2 + G^2)/[(N+1)(\beta_1 H + G)^2]}. \tag{67}$$

Since $N - 1 < N + 1$ and $(\beta_1^2 H^2 + G^2) < (\beta_1 H + G)^2$, the following inequality always holds:

$$\frac{N-1}{N+1} \cdot \frac{\beta_1^2 H^2 + G^2}{(\beta_1 H + G)^2} < 1. \tag{68}$$

Therefore, in the high-SNR scenario, we have

$$\frac{\beta_1^2 \mathcal{B}^2}{2N\mathcal{A} - \mathcal{C}^2} > \frac{3}{5}. \tag{69}$$

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," *Comput. Netw.*, vol. 38, no. 4, pp. 393–422, Mar. 2002.

[2] K. Romer and F. Mattern, "The design space of wireless sensor networks," *IEEE Wireless Commun.*, vol. 11, no. 6, pp. 54–61, Dec. 2004.

[3] N. Bulusu and S. Jha, *Wireless Sensor Networks: A Systems Perspective*. Norwood, MA: Artech House, 2005.

[4] W. Su, "Overview of time synchronization issues in sensor networks," in *Embedded Systems*, R. Zurawski, Ed. Boca Raton, FL: CRC, 2005.

[5] H. Kopetz and W. Ochsenreiter, "Clock synchronization in distributed real-time systems," *IEEE Trans. Comput.*, vol. C-36, no. 8, pp. 933–940, Aug. 1987.

[6] D. L. Mills, "Internet time synchronization: The network time protocol," *IEEE Trans. Commun.*, vol. 39, no. 10, pp. 1482–1493, Oct. 1991.

[7] S. Ganeriwal, R. Kumar, and M. B. Srivastava, "Timing-sync protocol for sensor networks," in *Proc. 1st Int. Conf. Embedded Netw. Sens. Syst.*, 2003, pp. 138–149.

[8] J. Elson, L. Girod, and D. Estrin, "Fine-grained network time synchronization using reference broadcasts," in *Proc. 5th Symp. Oper. Syst. Des. Implement.*, Boston, MA, Dec. 2002, pp. 147–163.

[9] K.-L. Noh, E. Serpedin, and K. Qaraqe, "A new approach for time synchronization in wireless sensor networks: Pairwise broadcast synchronization," *IEEE Trans. Wireless Commun.*, vol. 7, no. 9, pp. 3318–3322, Dec. 2008.

[10] M. Maroti, B. Kusy, G. Simon, and A. Ledeczi, "The flooding time synchronization protocol," in *Proc. 2nd Int. Conf. Embedded Netw. Sens. Syst.*, 2004, pp. 39–49.

[11] W. Su and I. F. Akyildiz, "Time-diffusion synchronization protocol for wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 13, no. 2, pp. 384–397, Apr. 2005.

[12] I. Sari, E. Serpedin, and B. Suter, "Application of Gibbs Sampler for clock synchronization in RBS-protocol," in *Proc. Military Commun. Conf.*, Oct. 2006, pp. 1–4.

[13] A. Syed and J. Heidemann, "Time synchronization for high latency acoustic networks," in *Proc. IEEE Infocom*, Barcelona, Spain, Apr. 2006, pp. 1–12.

[14] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[15] K.-L. Noh, Q. M. Chaudhari, E. Serpedin, and B. W. Suter, "Novel clock phase offset and skew estimation using two-way timing message exchanges for wireless sensor networks," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 766–777, Apr. 2007.
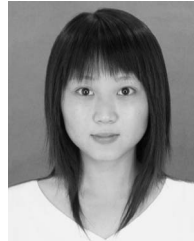
Mei Leng received the B.Eng. degree from The University of Electronic Science and Technology of China, Chengdu, China, in 2005. She is currently working toward the Ph.D. degree with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong.

Her current research interests include statistical signal processing, optimization, machine learning, and Bayesian analysis, with applications to wireless sensor networks and wireless communication systems.
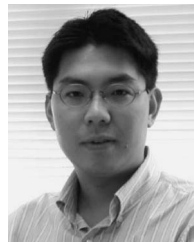
Yik-Chung Wu received the B.Eng. (EEE) and M.Phil. degrees from The University of Hong Kong (HKU), Hong Kong, in 1998 and 2001, respectively, and the Ph.D. degree from Texas A&M University, College Station, in 2005.

After receiving the Master's degree, he became a Research Assistant with HKU. While attending Texas A&M University, he was fully supported by a scholarship from the prestigious Croucher Foundation. From August 2005 to August 2006, he was with Thomson Corporate Research, Princeton, NJ, as a Member of Technical Staff. Since September 2006, he has been with the Department of Electrical and Electronic Engineering, HKU, as an Assistant Professor. His research interests include signal processing and communication systems, particularly receiver algorithm design, synchronization techniques, channel estimation, and equalization.

Dr. Wu was a Technical Program Committee member of the Fall 2005 IEEE Vehicular Technology Conference, Globecom in 2006 and 2008, and the IEEE International Conference on Communications in 2007 and 2008. He is currently serving as an Associate Editor for the IEEE COMMUNICATIONS LETTERS.