On Minimum Variance Unbiased Estimation of Clock Offset in a Two-Way Message Exchange Mechanism

Qasim M. Chaudhari, Erchin Serpedin, and Khalid Qaraqe, *Senior Member, IEEE*

Abstract—For many applications, distributed networks require the local clocks of the constituent nodes to run close to an agreed upon notion of time. Most of the widely used clock synchronization algorithms in such systems employ the sender-receiver protocol based on a two-way timing message exchange paradigm. Maximum likelihood estimator (MLE) of the clock offset based on the timing message exchanges between two clocks was derived in D. R. Jeske, On maximum likelihood estimation of clock offset[*IEEE Trans. Commun.*, vol. 53, pp. 53–54, Jan. 2005], when the fixed delays are symmetric and the variable delays in each direction assume an exponential distribution with an unknown mean. Herein, the best linear unbiased estimate using order statistics (BLUE-OS) of the clock offset between two nodes is derived assuming both symmetric and asymmetric exponential network delays, respectively. The Rao-Blackwell-Lehmann-Scheffé theorem is then exploited to obtain the minimum variance unbiased estimate (MVUE) for the clock offset which it is shown to coincide with the BLUE-OS. In addition, it is found that the MVUE of the clock offset in the presence of symmetric network delays also coincides with the MLE. Finally, in the presence of asymmetric network delays, although the MLE is biased, it is shown to achieve lesser mean-square error (MSE) than the MVUE in the region around the point where the bidirectional network link delays are symmetric and hence its merit as the most versatile estimator is fairly justified.

Index Terms—Clock, estimation, signal processing, synchronization, wireless sensor networks.

I. INTRODUCTION

IN distributed systems, maintaining the logical clocks of the computers in such a way that they are never too far apart is one of the most challenging problems. Whether it is the disciplining of computer clocks with the devices synchronized to a GPS satellite or a network time protocol (NTP) server over the Internet, it is possible to equip some primary time servers for the purpose of synchronizing a much larger number of secondary servers and clients connected through a common infrastructure. In order to do this, a distributed network clock synchronization protocol is required through which a server clock can be read, the readings to other clients can be transmitted and each client clock can be adjusted as required. In such a distributed synchronization approach, the participating devices exchange timing information with their chosen reference at regular intervals and adjust their logical clocks accordingly.

A computer clock in general has two components, namely a frequency source and a means of accumulating timing events (consisting of a clock interrupt mechanism and a counter implemented in software). The implementation of the computer clock in the operating system and the programming interface differ between operating systems and hardware platforms. However, the basic source of timing offsets are an uncompensated quartz crystal oscillator and the clock interrupts it generates. Theoretically, two clocks would remain synchronized if their offsets are set equal and their frequency sources run at the same rate. However, practical clocks are set with limited precision and the frequency sources run at slightly different rates. In addition, the frequency of a crystal oscillator varies due to initial manufacturing tolerance, aging, temperature, pressure and other factors. Because of these inherent instabilities, distributed clocks must regularly be synchronized to keep them running close to each other.

Clock synchronization is important for many applications such as Internet delay measurements, cellular networks, data security algorithms, MAC protocols like time division multiple access (TDMA), IP telephony, ordering of updates in database systems, etc. Recently, with the advent of wireless sensor networks (WSNs), developing clock synchronization protocols that suit their specific requirements is becoming an important research problem [2]. A large number of WSN applications require the clocks of the nodes to run synchronously on a common timescale. This is the case with applications such as data fusion, efficient duty cycling operations, acoustic beamforming, localization, security, and object tracking. Unlike conventional networks, energy efficiency must also be taken into account for addressing the clock synchronization problem in WSNs.

II. RELATED WORK

During the last two decades, many clock synchronization protocols have been proposed such as [3], [5], [6], etc. The NTP [3] is a protocol for synchronizing the clocks of computer systems over packet-switched, variable-latency data networks and it represents the Internet standard for time synchronization. It is a layered client-server architecture based on the UDP message passing which synchronizes computer clocks in a hierarchical way using the offset delay estimation method. NTP's sender-receiver synchronization mechanism is widely accepted in designing time synchronization algorithms and consists of the

same two-way timing message exchange paradigm targeted in this paper.

A protocol based on the remote clock reading method was put forward by [5], which handles unbounded message delays between processes. In [6], the time transmission protocol is used by a node to communicate the time on its clock to a target node, which subsequently estimates the time in the source node by using message timestamps and message delay statistics.

For *ad-hoc* communication networks, the time synchronization protocol [7] represented one of the pioneering contributions reported in this area. The protocol is based on generating timestamps to record the time at which an event of interest occurs. The timestamps are updated by each node using its local clock and the time transformation method, where the final timestamp is expressed in terms of an interval with a lower bound and an upper bound.

In the realm of WSNs, several clock synchronization protocols of special interest were proposed: reference broadcast synchronization (RBS [8]), timing synch protocol for sensor networks (TPSN [4]) and time diffusion protocol (TDP [9]). RBS relies on the simultaneous reception of broadcast pulses by several nodes transmitted by a common neighboring node after which the nodes exchange their timestamps and estimate the relative time offsets and skews. On the other hand, TPSN is based on the same sender-receiver paradigm as in NTP, like many other traditional clock synchronization protocols. The basic difference is that TPSN has been molded sufficiently well to suit the WSN requirements. TDP establishes a network-wide equilibrium time through an iterative weighted averaging technique by diffusing the synchronization messages through all the network nodes. In the context of WSNs, the results presented in this paper are applicable to a wide range of time transfer problems either directly (as is the case with TPSN and other protocols) or through some minor extensions.

Based on the two-way timing message exchange mechanism and assuming an exponential network delay distribution, [10] brought some interesting results by applying techniques from the statistical signal processing field. However, an important problem was not fully addressed in [10]. Assuming known fixed delays and known symmetric exponential mean link delays, [10] concluded that the maximum likelihood estimator (MLE) of the clock offset does not exist in this scenario. However, in [1], assuming an unknown fixed delay, irrespective of the symmetric exponential mean link delay being known or unknown, the MLE of the clock offset was successfully derived.

Assuming both symmetric and asymmetric exponentially distributed network link delays, this paper focuses on finding the best linear unbiased estimators using order statistics (BLUE-OS) and the minimum variance unbiased estimates (MVUE) for the clock offset between two nodes and evaluates their performance in terms of the mean square error (MSE), which is chosen as the performance criterion throughout this paper. The timing exchange mechanism between the two nodes is the same classical two-way message exchange mechanism adopted in protocols such as NTP [3], TPSN [4], etc. The main contributions of this paper are as follows.

1) A relatively unnoticed estimation scheme in engineering literature, the BLUE-OS, is investigated in the context
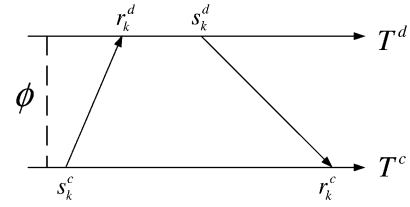

Fig. 1. Sender-receiver timing message exchange paradigm.

of clock offset and relevant clock offset estimators are derived.

2) The Rao-Blackwell-Lehmann-Scheffé theorem is then used to derive the MVUE and it is shown that the MVUE coincides with the BLUE-OS. Therefore, in the class of unbiased estimators, it is the optimal solution and no other estimator can be found with lesser MSE (or variance, which is the same as MSE in the unbiased case) than the MVUE. For the sake of completion, the clock offset estimators are also derived in two scenarios, namely when the mean of the exponential link delays is known and unknown for each direction, respectively.

3) A short commentary on whether the MVUE is the best possible solution as compared to the other estimators such as the MLE is presented. It is shown that in the most practical scenario, i.e., asymmetric link delays with unknown exponential means, the MLE derived in the presence of symmetric link delays—although biased for asymmetric link delays—outperforms the MVUE in terms of the achievable MSE in the region around the point of link symmetry.

III. Signaling Mechanism

Adopting the classical approach of sender-receiver synchronization for performing a timing handshake between a pair of nodes, the uplink and downlink timing message exchanges between two clocks c and d are shown in Fig. 1. The messages s_k^c and r_k^c represent the times measured by the local clock of node c, while the messages r_k^d and s_k^d represent the times measured by the local clock of node d (which is also the reference). The synchronization procedure starts at time s_1^c and at each successive message exchange round k, node c sends a synchronization packet containing the timestamp s_k^c to node d which records its reception time as r_k^d. At s_k^d, node d sends an acknowledgement packet back to node c containing the timestamps r_k^d and s_k^d, which is delivered and timestamped at time r_k^c in accordance with node c clock. This process between the two nodes is repeated N times, where N stands for the required number of samples. It should be noted that N is a function of the target synchronization accuracy and the price the protocol is willing to invest in the form of network resources.

Based on the above pairwise synchronization message exchange mechanism, the clock offset measurement model can be represented in terms of these two equations

$$r_k^d = s_k^c + \delta + \phi + X_k$$
$$r_k^c = s_k^d + \delta - \phi + Y_k.$$

For simplification, the above equations will be rewritten as

$$U_k = \delta + \phi + X_k,$$
$$V_k = \delta - \phi + Y_k$$

where $U_k \triangleq r_k^d - s_k^c$ and $V_k \triangleq r_k^c - s_k^d$. The quantity δ symbolizes the fixed portions of the delays assumed to be symmetric for each direction, X_k and Y_k denote the independent and identically distributed variable portions of delays and assume exponential distributions with means α and β, respectively, and ϕ stands for the clock offset of node d with respect to node c. Turning towards modeling the network delays, the Weibull, exponential, Gamma, and log-normal distributions [11]–[13] have been found to closely capture the true distribution of the network delays. There are various reasons behind choosing the exponential distribution for the purpose of this study. Reference [14] collected several traces of delay measurements on the Internet and MBone [15] for more than a month using constant length UDP packets whose payloads consisted of a sequence number and a timestamp sent out at periodic intervals. The exponential distribution provided quite a satisfactory fit for the measurements obtained in the experiment. In addition, a single-server M/M/1 queue can fittingly represent the cumulative link delay for point-to-point hypothetical reference connections, where the random delays are independently modeled as exponential random variables [10]. Moreover, [10] proposed five different clock offset estimation algorithms such as the median round delay, the minimum round delay, the minimum link delay, the median phase and the average phase, amongst which the minimum link delay algorithm has been experimentally demonstrated to be superior than the rest [16]. Reference [1] later mathematically proved that this algorithm yields the maximum likelihood estimate under exponential link delays. All these results confirm that the assumption of exponential distribution for network delays is a sufficiently adequate model to fit the experimental observations.

In [1], it was argued that for an unknown δ, irrespective of the symmetric exponential distribution mean λ (which is the same as α and β in symmetric case) being known or unknown, the MLE of the vector parameter $\mathbf{\Phi_{MLE}^S} = [\delta \phi \lambda]$ is given by

$$\mathbf{\Phi_{MLE}^S} = \begin{bmatrix} \hat{\delta}_{MLE}^S \\ \hat{\phi}_{MLE}^S \\ \hat{\lambda}_{MLE}^S \end{bmatrix} = \frac{1}{2} \begin{bmatrix} U_{(1)} + V_{(1)} \\ U_{(1)} - V_{(1)} \\ \bar{U} + \bar{V} - (U_{(1)} + V_{(1)}) \end{bmatrix} \tag{1}$$

where $U_{(1)}$ and $V_{(1)}$ denote the minimum order statistics (i.e., $U_{(1)} = \min_k U_k$), and $\bar{U}$ and $\bar{V}$ represent the sample average of the data $\{U_k\}_{k=1}^N$ and $\{V_k\}_{k=1}^N$, respectively. When λ is known, the MLEs of $\{\delta, \phi\}$ assume the same expressions.

Next, the BLUE-OS and MVUE are derived for both the asymmetric and symmetric cases, assuming unknown exponential delay means. For the sake of completion, we also state the results for the known case at the end of the corresponding subsections of Section V.

IV. BLUE-OS

Deriving regular BLUE for a problem yields suboptimal results in general, since the class of unbiased estimators, within which the search is performed, is restricted to be linear. In the case when the noise is normally distributed, the direct application of BLUE provides the optimal solution by virtue of the Gauss–Markov theorem. But for other distributions, including the exponential distribution as is the case with the modeling framework adopted in this paper, the direct application of BLUE is not of much relevance. However, for a general location-scale distribution, [17] suggested a different technique based on the derivation of BLUE using order statistics instead of just the raw observations. Such a technique will be applied herein to the target scenario as follows.

Let the order statistics of the observations $\{U_k\}_{k=1}^N$ and $\{V_k\}_{k=1}^N$ be denoted as $\{U_{(k)}\}_{k=1}^N$ and $\{V_{(k)}\}_{k=1}^N$, respectively. Define

$$U_k' \triangleq \frac{1}{\alpha}(U_k - \delta - \phi)$$
$$V_k' \triangleq \frac{1}{\beta}(V_k - \delta + \phi)$$

which are a set of independent observations on the standardized variate and hence their distribution is parameter-free. The order statistics of U_k' and V_k' are denoted by $U_{(k)}'$ and $V_{(k)}'$, respectively. The following relations hold:

$$E\left[U_{(k)}\right] = \delta + \phi + \alpha E\left[U_{(k)}'\right]$$
$$E\left[V_{(k)}\right] = \delta - \phi + \beta E\left[V_{(k)}'\right]$$
$$\text{var}\left[U_{(k)}\right] = \alpha^2 \text{var}\left[U_{(k)}'\right]$$
$$\text{var}\left[V_{(k)}\right] = \beta^2 \text{var}\left[V_{(k)}'\right]$$
$$\text{cov}\left[U_{(k)}U_{(j)}\right] = \alpha^2 \text{cov}\left[U_{(k)}'U_{(j)}'\right]$$
$$\text{cov}\left[V_{(k)}V_{(j)}\right] = \beta^2 \text{cov}\left[V_{(k)}'V_{(j)}'\right].$$

Now using standard results from [18, p. 500], the statistics of the ordered samples can be expressed as

$$E\left[U_{(k)}'\right] = E\left[V_{(k)}'\right] = \sum_{i=1}^k \frac{1}{(N-i+1)}$$
$$\text{var}\left[U_{(k)}'\right] = \text{var}\left[V_{(k)}'\right] = \sum_{i=1}^k \frac{1}{(N-i+1)^2}$$
$$\text{cov}\left[U_{(k)}'U_{(j)}'\right] = \text{cov}\left[V_{(k)}'V_{(j)}'\right] = \sum_{i=1}^k \frac{1}{(N-i+1)^2}.$$

As a result, the $N \times N$ symmetric positive-definite covariance matrix $\mathbf{C}$ for both $U_{(k)}'$ and $V_{(k)}'$ takes the common expression:

$$\mathbf{C} = \begin{bmatrix} \frac{1}{N^2} & \frac{1}{N^2} & \cdots & \frac{1}{N^2} \\ \frac{1}{N^2} & \frac{1}{N^2} + \frac{1}{(N-1)^2} & \cdots & \frac{1}{N^2} + \frac{1}{(N-1)^2} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{1}{N^2} & \frac{1}{N^2} + \frac{1}{(N-1)^2} & \cdots & \sum_{k=1}^N \frac{1}{(N-k+1)^2} \end{bmatrix}.$$

It is worth pointing out at this time that a simple exercise yields the following closed-form expression for the inverse of the covariance matrix (see the first matrix at the bottom of the page).

The BLUE-OS will be next derived separately for both symmetric and asymmetric network delays.

A. Symmetric Link Delays

The symmetric network delay assumption is reasonable for some realistic scenarios, specially when the nodes have a direct communication link between them and the topology of the network is constant. In this case, $\alpha = \beta \triangleq \lambda$. Consider the BLUE-OS $\boldsymbol{\Phi}_{\mathbf{BLUE-OS}}^{\mathbf{S}} \triangleq [\delta\phi\lambda]^T$, which is a linear function of an ordered set of observations $\{U_{(k)}\}_{k=1}^N$ and $\{V_{(k)}\}_{k=1}^N$. Let $\mathbf{z} \triangleq [U_{(1)}U_{(2)}\ldots U_{(N)}V_{(1)}V_{(2)}\ldots V_{(N)}]^T$. Then, it is straightforward to notice that (see the second matrix at the bottom of the page), where $\mathbf{z}$ is the $2N \times 1$ ordered data vector, $\mathbf{Q}$ is a known matrix of dimension $2N \times 3$ and $\boldsymbol{\Phi}_{\mathbf{BLUE-OS}}^{\mathbf{S}}$ is the 3×1 vector of unknown parameters. The above linear relationship lends the problem to be solved by the Gauss–Markov theorem. Since $\{U_{(k)}\}_{k=1}^N$ and $\{V_{(k)}\}_{k=1}^N$ are independent data sets

as shown in the third matrix at the bottom of the page. Therefore, the BLUE-OS in the symmetric exponential network delays case is given by (2), shown at the bottom of the page, with $\bar{U}$ and $\bar{V}$ representing the sample averages of the data sets $\{U_k\}_{k=1}^N$ and $\{V_k\}_{k=1}^N$, respectively, and which coincide with the sample averages of ordered observations $\{U_{(k)}\}_{k=1}^N$ and $\{V_{(k)}\}_{k=1}^N$, respectively. Note that the BLUE-OS of the clock offset matches the MLE in (1), but this is not true for the estimates of the deterministic and mean variable delay (δ and λ). This is because $\hat{\phi}_{\mathrm{MLE}}^S$ is unbiased unlike $\hat{\delta}_{\mathrm{MLE}}^S$ and $\hat{\lambda}_{\mathrm{MLE}}^S$, which are multiplied by suitable factors to remove their bias.

B. Asymmetric Link Delays

In many broadband and wireless channels, and *ad-hoc* networks with time-varying topologies, the symmetric network delay assumption does not hold and applying the same results derived under the symmetric assumption is suboptimal. Therefore, a study for deriving the efficient estimators in this case is very important. Let $\boldsymbol{\Phi}_{\mathbf{BLUE-OS}}^{\mathbf{A}} \triangleq [\delta\phi\alpha\beta]^T$, then the linear model based on the ordered observations can be expressed as

$$\mathbf{C}^{-1} = \begin{bmatrix} N^2 + (N-1)^2 & -(N-1)^2 & 0 & 0 & \cdots & 0 \\ -(N-1)^2 & (N-1)^2 + (N-2)^2 & -(N-2)^2 & 0 & \cdots & 0 \\ 0 & -(N-2)^2 & (N-2)^2 + (N-3)^2 & -(N-3)^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$E[\mathbf{z}] = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 & -1 & -1 & \cdots & -1 \\ \frac{1}{N} & \frac{1}{N} + \frac{1}{N-1} & \cdots & \sum_{k=1}^N \frac{1}{(N-k+1)} & \frac{1}{N} & \frac{1}{N} + \frac{1}{N-1} & \cdots & \sum_{k=1}^N \frac{1}{(N-k+1)} \end{bmatrix}^T \begin{bmatrix} \delta \\ \phi \\ \lambda \end{bmatrix}$$
$$= \mathbf{Q}\boldsymbol{\Phi}_{\mathbf{BLUE-OS}}^{\mathbf{S}}$$

$$\left(\mathbf{Q}^T\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{Q}\right)^{-1}\mathbf{Q}^T\mathbf{C}_{\mathbf{z}}^{-1} = \frac{1}{2N(N-1)} \begin{bmatrix} N^2 - 1 & -1 & \cdots & -1 & N^2 - 1 & -1 & \cdots & -1 \\ N(N-1) & 0 & \cdots & 0 & -N(N-1) & 0 & \cdots & 0 \\ N - N^2 & N & \cdots & N & N - N^2 & N & \cdots & N \end{bmatrix}$$

$$\begin{aligned} \hat{\boldsymbol{\Phi}}_{\mathbf{BLUE-OS}}^{\mathbf{S}} = \left(\mathbf{Q}^T\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{Q}\right)^{-1}\mathbf{Q}^T\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{z} &= \begin{bmatrix} \hat{\delta}_{BLUE-OS}^S \\ \hat{\phi}_{BLUE-OS}^S \\ \hat{\lambda}_{BLUE-OS}^S \end{bmatrix} \\ &= \frac{1}{2N(N-1)} \begin{bmatrix} (N^2-1)U_{(1)} - \sum_{k=2}^N U_{(k)} + (N^2-1)V_{(1)} - \sum_{k=2}^N V_{(k)} \\ N(N-1)U_{(1)} - N(N-1)V_{(1)} \\ (N-N^2)U_{(1)} + \sum_{k=2}^N U_{(k)} + (N-N^2)V_{(1)} + \sum_{k=2}^N V_{(k)} \end{bmatrix} \\ &= \frac{1}{2(N-1)} \begin{bmatrix} N\left(U_{(1)} + V_{(1)}\right) - \left(\bar{U} + \bar{V}\right) \\ (N-1)\left(U_{(1)} - V_{(1)}\right) \\ N\left\{\left(\bar{U} + \bar{V}\right) - \left(U_{(1)} + V_{(1)}\right)\right\} \end{bmatrix} \quad (2) \end{aligned}$$

the first equation at the bottom of the page, where $\mathbf{z}$ is again a $2N \times 1$ concatenated vector of ordered data $U_{(k)}$ and $V_{(k)}$, $\mathbf{Q}$ is a known matrix of dimension $2N \times 4$ and $\mathbf{\Phi}_{\mathbf{BLUE-OS}}^{\mathbf{A}}$ is the 4×1 vector of unknown parameters. The inverse of the joint covariance matrix $\mathbf{C_z}$ for $U_{(k)}$ and $V_{(k)}$ can be expressed as

$$\mathbf{C_z}^{-1} = \frac{1}{\alpha^2 \beta^2} \begin{bmatrix} \beta^2 \mathbf{C}^{-1} & \mathbf{0} \\ \mathbf{0} & \alpha^2 \mathbf{C}^{-1} \end{bmatrix}.$$

Based on the above expression, it follows that

$$\left(\mathbf{Q}^T \mathbf{C_z}^{-1} \mathbf{Q}\right)^{-1} = \frac{1}{2N(N-1)}$$
$$\times \begin{bmatrix} \frac{1}{2}(\alpha^2 + \beta^2) & \frac{1}{2}(\alpha^2 - \beta^2) & -\alpha^2 & -\beta^2 \\ \frac{1}{2}(\alpha^2 - \beta^2) & \frac{1}{2}(\alpha^2 + \beta^2) & -\alpha^2 & \beta^2 \\ -\alpha^2 & -\alpha^2 & 2N\alpha^2 & 0 \\ -\beta^2 & \beta^2 & 0 & 2N\beta^2 \end{bmatrix}$$

and consequently (see the second equation at the bottom of the page), which implies (3), shown at the bottom of the page.

V. MVUE

In parameter estimation, very often the ultimate goal is to find the estimator that achieves the minimum MSE (MMSE), and it is usually the criterion of choice. However, it is well known in theory that the optimal MSE estimators are usually not realizable. Since the MSE is the sum of estimator variance and squared bias, a technique chosen to attain realizable yet best estimators is to constrain the bias to be zero (since the dependence of MMSE estimator on the unknown parameter typically comes from the bias). Therefore, restricting the possible estimators to be unbiased and then finding the estimator with the smallest variance for all values of the unknown parameter yields the optimal solution within the class of unbiased estimators. Therefore, we will resort on the concept of MVUE, which is derived based on the Rao-Blackwell-Lehmann-Scheffé theorem.

A. Asymmetric Link Delays

Starting with the asymmetric case, the likelihood function for the clock offset as a function of observations $\{U_k\}_{k=1}^N$ and $\{V_k\}_{k=1}^N$ is given by

$$L(\delta, \phi, \alpha, \beta)$$
$$= \alpha^{-N} \exp\left[-\frac{1}{\alpha} \sum_{k=1}^N \{U_k - \delta - \phi\}\right]$$
$$\cdot \beta^{-N} \exp\left[-\frac{1}{\beta} \sum_{k=1}^N \{V_k - \delta + \phi\}\right]$$
$$\cdot u\left[U_{(1)} - \delta - \phi\right] \cdot u\left[V_{(1)} - \delta + \phi\right] \quad (4)$$

$$E[\mathbf{z}] = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 & -1 & -1 & \cdots & -1 \\ \frac{1}{N} & \frac{1}{N} + \frac{1}{N-1} & \cdots & \sum_{k=1}^N \frac{1}{(N-k+1)} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \frac{1}{N} & \frac{1}{N} + \frac{1}{N-1} & \cdots & \sum_{k=1}^N \frac{1}{(N-k+1)} \end{bmatrix}^T \begin{bmatrix} \delta \\ \phi \\ \alpha \\ \beta \end{bmatrix}$$
$$= \mathbf{Q}\mathbf{\Phi}_{\mathbf{BLUE-OS}}^{\mathbf{A}}$$

$$\left(\mathbf{Q}^T \mathbf{C_z}^{-1} \mathbf{Q}\right)^{-1} \mathbf{Q}^T \mathbf{C_z}^{-1} = \frac{1}{2N(N-1)} \begin{bmatrix} N^2-1 & -1 & \cdots & -1 & N^2-1 & -1 & \cdots & -1 \\ N^2-1 & -1 & \cdots & -1 & -(N^2-1) & 1 & \cdots & 1 \\ 2N-2N^2 & 2N & \cdots & 2N & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 2N-2N^2 & 2N & \cdots & 2N \end{bmatrix}$$

$$\hat{\mathbf{\Phi}}_{\mathbf{BLUE-OS}}^{\mathbf{A}} = \left(\mathbf{Q}^T \mathbf{C_z}^{-1} \mathbf{Q}\right)^{-1} \mathbf{Q}^T \mathbf{C_z}^{-1} \mathbf{z} = \begin{bmatrix} \hat{\delta}_{BLUE-OS}^A \\ \hat{\phi}_{BLUE-OS}^A \\ \hat{\alpha}_{BLUE-OS}^A \\ \hat{\beta}_{BLUE-OS}^A \end{bmatrix}$$
$$= \frac{1}{2N(N-1)} \begin{bmatrix} (N^2-1)U_{(1)} - \sum_{k=2}^N U_{(k)} + (N^2-1)V_{(1)} - \sum_{k=2}^N V_{(k)} \\ (N^2-1)U_{(1)} - \sum_{k=2}^N U_{(k)} - (N^2-1)V_{(1)} + \sum_{k=2}^N V_{(k)} \\ 2NU_{(1)} - 2N^2 U_{(1)} + 2N\sum_{k=2}^N U_{(k)} \\ 2NV_{(1)} - 2N^2 V_{(1)} + 2N\sum_{k=2}^N V_{(k)} \end{bmatrix}$$
$$= \frac{1}{2(N-1)} \begin{bmatrix} N\left(U_{(1)} + V_{(1)}\right) - (\bar{U} + \bar{V}) \\ N\left(U_{(1)} - V_{(1)}\right) - (\bar{U} - \bar{V}) \\ 2N\left(\bar{U} - U_{(1)}\right) \\ 2N\left(\bar{V} - V_{(1)}\right) \end{bmatrix}. \quad (3)$$

where $u[\cdot]$ denotes the unit step function. Exploiting the fact that the raw sample mean and the ordered sample mean are actually the same, (4) can be factored as

$$
\begin{aligned}
L(\delta, &\phi, \alpha, \beta) \\
&= g_1\left(\sum_{k=1}^{N} U_{(k)}, \delta, \phi, \alpha\right) \\
&\quad \times g_2\left(\sum_{k=1}^{N} V_{(k)}, \delta, \phi, \beta\right) \\
&\quad \times g_3\left(U_{(1)}, \delta, \phi\right) \\
&\quad \times g_4(V_{(1)}, \delta, \phi) h_1(U_k, V_k)
\end{aligned}
$$

where

$$
\begin{aligned}
g_1\left(\sum_{k=1}^{N} U_{(k)}, \delta, \phi, \alpha\right) &= \alpha^{-N} e^{-\frac{1}{\alpha}\sum_{k=1}^{N}(U_{(k)}-\delta-\phi)} \\
g_2\left(\sum_{k=1}^{N} V_{(k)}, \delta, \phi, \beta\right) &= \beta^{-N} e^{-\frac{1}{\beta}\sum_{k=1}^{N}(V_{(k)}-\delta+\phi)} \\
g_3\left(U_{(1)}, \delta, \phi\right) &= u\left[U_{(1)} - \delta - \phi\right] \\
g_4\left(V_{(1)}, \delta, \phi\right) &= u\left[V_{(1)} - \delta + \phi\right], h_1(U_k, V_k) = 1.
\end{aligned}
$$

In the above relations, $h_1(U_k, V_k)$ is independent of the unknown vector parameter $\mathbf{\Phi}_{\mathbf{MVUE}}^{\mathbf{A-U}} = [\delta \phi \alpha \beta]^T$, whereas $g_1(\sum_{k=1}^{N} U_{(k)}, \delta, \phi, \alpha)$, $g_2(\sum_{k=1}^{N} V_{(k)}, \delta, \phi, \beta)$, $g_3(U_{(1)}, \delta, \phi)$ and $g_4(V_{(1)}, \delta, \phi)$ are functions depending on the data through $\mathbf{T} = \{\sum_{k=1}^{N} U_{(k)}, U_{(1)}, \sum_{k=1}^{N} V_{(k)}, V_{(1)}\}$. Therefore, according to Neyman-Fisher factorization theorem, $\mathbf{T}$ is a sufficient statistic for $\mathbf{\Phi}_{\mathbf{MVUE}}^{\mathbf{A-U}}$.

Since $\dim(\mathbf{T}) = \dim(\mathbf{\Phi}_{\mathbf{MVUE}}^{\mathbf{A-U}})$, it is easier to determine the MVUE directly from $\mathbf{T}$ without having to evaluate $E[\check{\mathbf{\Phi}}_{\mathbf{MVUE}}^{\mathbf{A-U}} | \mathbf{T}]$ by finding a 4×1 vector function $\hat{\mathbf{\Phi}}_{\mathbf{MVUE}}^{\mathbf{A-U}}$ such that $E[\hat{\mathbf{\Phi}}_{\mathbf{MVUE}}^{\mathbf{A-U}}] = \mathbf{\Phi}_{\mathbf{MVUE}}^{\mathbf{A-U}}$, provided that $\mathbf{T}$ is a complete sufficient statistic. Finding the probability density function (pdf) of $\mathbf{T}$ is required to prove that $\mathbf{T}$ is complete but the problem of finding this pdf is a little complex, because $\sum_{k=1}^{N} U_{(k)}$ and $U_{(1)}$, and similarly $\sum_{k=1}^{N} V_{(k)}$ and $V_{(1)}$, are not independent.

The joint pdf of $U_{(1)}, U_{(2)}, \ldots, U_{(N)}$ is given by

$$
\begin{aligned}
p\left(U_{(1)}, U_{(2)}, \ldots, U_{(N)}\right) &= N! \alpha^{-N} e^{-\frac{1}{\alpha}\sum_{k=1}^{N}\{U_{(k)}-\delta-\phi\}} \\
&\quad \cdot \prod_{k=1}^{N} u\left[U_{(k)} - \delta - \phi\right] \quad (5)
\end{aligned}
$$

whereas the pdf of the minimum order statistic $U_{(1)}$ is also exponential with mean α/N. Now consider the transformation

$$
z_k = (N-k+1)\left(U_{(k)} - U_{(k-1)}\right), \\
k = 1, 2, \ldots, N
$$

where $U_{(0)} = \delta + \phi$. Since $\sum_{k=1}^{N}(U_{(k)} - \delta - \phi) = \sum_{k=1}^{N} z_k$ and the Jacobian of the transformation is $N!$, a substitution in (5) reveals that

$$
p(z_1, z_2, \ldots, z_N) = \alpha^{-N} e^{-\frac{1}{\alpha}\sum_{k=1}^{N} z_k} \cdot \prod_{k=1}^{N} u[z_k],
$$

i.e., z_k are *independent* exponential random variables with similar mean α. In addition, since each $z_k \sim \exp(\alpha)$, each z_k assumes a Gamma distribution $z_k \sim \Gamma(1, \alpha)$, too. Using the relationship $\sum_{k=1}^{N}(U_{(k)} - U_{(1)}) = \sum_{k=2}^{N} z_k$, and the fact that each of $z_2, z_3, \ldots, z_N$ is independent of z_1 (and hence of $U_{(1)}$, since $z_1 = N(U_{(1)} - \delta - \phi)$), $\sum_{k=1}^{N}(U_{(k)} - U_{(1)}) \sim \Gamma(N-1, \alpha)$ and is independent of $U_{(1)}$.

By a similar reasoning, it can be deduced that $\sum_{k=1}^{N}(V_{(k)} - V_{(1)}) \sim \Gamma(N-1, \beta)$ and is independent of $V_{(1)}$. Therefore, the one-to-one function $\mathbf{T}' = \{\sum_{k=1}^{N}(U_{(k)} - U_{(1)}), U_{(1)}, \sum_{k=1}^{N}(V_{(k)} - V_{(1)}), V_{(1)}\}$ of $\mathbf{T}$ is also sufficient for estimating $\mathbf{\Phi}_{\mathbf{MVUE}}^{\mathbf{A-U}}$ because the sufficient statistics are unique within one-to-one transformations [19]. Consequently, $\mathbf{T}'$ comprises of four independent random variables that in terms of the three-parameter Gamma distribution assume the distributions:

$$
\begin{aligned}
r &= \sum_{k=1}^{N}(U_{(k)} - U_{(1)}) \sim \Gamma(N-1, \alpha, 0) \\
s &= \sum_{k=1}^{N}(V_{(k)} - V_{(1)}) \sim \Gamma(N-1, \beta, 0) \\
U_{(1)} &\sim \Gamma(1, \alpha/N, \delta + \phi) \\
V_{(1)} &\sim \Gamma(1, \beta/N, \delta - \phi).
\end{aligned}
$$

Note that the domains of r and s are controlled by $U_{(1)}$ and $V_{(1)}$, respectively. Next, it has to be checked whether $\mathbf{T}'$, or equivalently $\mathbf{T}$, is complete. Completeness implies that there is but one function of $\mathbf{T}$ that is unbiased. Let $g(\mathbf{T}')$ be a function of $\mathbf{T}'$ such that $E[g(\mathbf{T}')] = \mathbf{\Phi}_{\mathbf{MVUE}}^{\mathbf{A-U}}$. Suppose that there exists another function h for which $E[h(\mathbf{T}')] = \mathbf{\Phi}_{\mathbf{MVUE}}^{\mathbf{A-U}}$ is also true. Then

$$
E[g(\mathbf{T}') - h(\mathbf{T}')] = E[\pi(\mathbf{T}')] = 0 \\
\forall \mathbf{\Phi}_{\mathbf{MVUE}}^{\mathbf{A-U}}
$$

where $\pi(\mathbf{T}') \triangleq g(\mathbf{T}') - h(\mathbf{T}')$ and the expectation is taken with respect to $p_{\mathbf{T}'}(\mathbf{T}'; \mathbf{\Phi}_{\mathbf{MVUE}}^{\mathbf{A-U}})$. As a result

$$
\begin{aligned}
&\iiiint_{R_{U_{(1)}, V_{(1)}}} \pi\left(r, U_{(1)}, s, V_{(1)}\right) \\
&\quad \cdot \frac{\alpha^{-(N-1)}}{\Gamma(N-1)} r^{N-2} e^{-\frac{r}{\alpha}} \cdot \frac{N}{\alpha} e^{-\frac{N}{\alpha}\{U_{(1)}-\delta-\phi\}} \\
&\quad \cdot \frac{\beta^{-(N-1)}}{\Gamma(N-1)} s^{N-2} e^{-\frac{s}{\beta}} \\
&\quad \cdot \frac{N}{\beta} e^{-\frac{N}{\beta}\{V_{(1)}-\delta+\phi\}} \, dr \, dU_{(1)} \, ds \, dV_{(1)} = 0 \\
&\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad \forall \mathbf{\Phi}_{\mathbf{MVUE}}^{\mathbf{A-U}}
\end{aligned}
$$

where $R_{U_{(1)}, V_{(1)}}$ is the region defined by $u[U_{(1)} - \delta - \phi]$ and $u[V_{(1)} - \delta - \phi]$. The above relation can be expressed as

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \pi\left(r, U_{(1)}, s, V_{(1)}\right)$$
$$\cdot r^{N-2} s^{N-2}$$
$$\times e^{-\left\{\frac{r}{\alpha} + \frac{N U_{(1)}}{\alpha} + \frac{s}{\beta} + \frac{N V_{(1)}}{\beta}\right\}} dr \, dU_{(1)} \, ds \, dV_{(1)} = 0$$
$$\forall \boldsymbol{\Phi}_{\mathbf{MVUE}}^{\mathbf{A-U}}.$$

The expression on the left above is the four-dimensional Laplace transform of the function $\pi(\mathbf{T}')$. It follows from the uniqueness theorem for two-sided Laplace transform that $\pi(\mathbf{T}') = 0$ almost everywhere, resulting in $g(\mathbf{T}') = h(\mathbf{T}')$ and hence there is only one unbiased function of $\mathbf{T}'$. This proves that the statistic $\mathbf{T}'$, or equivalently $\mathbf{T}$, is complete for estimating $\boldsymbol{\Phi}_{\mathbf{MVUE}}^{\mathbf{A-U}}$ when the links are asymmetric and both α and β are unknown.

Finally, the complete sufficient statistic $\mathbf{T}$ is also minimal owing to Bahadur's theorem which states that *if* $\mathbf{T}$, *taking values in* $\Re^k$, *is sufficient for* $\boldsymbol{\Phi}_{\mathbf{MVUE}}^{\mathbf{A-U}}$ *and boundedly complete, then* $\mathbf{T}$ *is minimal sufficient.*

What it remains to prove is finding an unbiased estimator for $\boldsymbol{\Phi}_{\mathbf{MVUE}}^{\mathbf{A-U}}$ as a function of $\mathbf{T}$, which will represent the MVUE according to the Rao-Blackwell-Lehmann-Scheffé theorem. Apparently, it seems difficult to find four unbiased functions of $\mathbf{T}$ for each of δ, ϕ, α and β just by inspection. But note that BLUE-OS $\hat{\boldsymbol{\Phi}}_{BLUE-OS}^{A}$ in (3) is also an unbiased function of $\mathbf{T}$. Hence, it is concluded that the BLUE-OS is also the MVUE

$$\hat{\boldsymbol{\Phi}}_{\mathbf{MVUE}}^{\mathbf{A-U}} = \begin{bmatrix} \hat{\delta}_{\mathrm{MVUE}}^{A-U} \\ \hat{\phi}_{\mathrm{MVUE}}^{A-U} \\ \hat{\alpha}_{\mathrm{MVUE}}^{A-U} \\ \hat{\beta}_{\mathrm{MVUE}}^{A-U} \end{bmatrix}$$
$$= \frac{1}{2(N-1)} \begin{bmatrix} N\left(U_{(1)} + V_{(1)}\right) - (\bar{U} + \bar{V}) \\ N\left(U_{(1)} - V_{(1)}\right) - (\bar{U} - \bar{V}) \\ 2N\left(\bar{U} - U_{(1)}\right) \\ 2N\left(\bar{V} - V_{(1)}\right) \end{bmatrix}. \quad (6)$$

As a result, the MVUE for the desired parameter, the clock offset, for asymmetric unknown network delays is expressed as

$$\hat{\phi}_{\mathrm{MVUE}}^{A-U} = \frac{1}{N-1} \left[N \frac{U_{(1)} - V_{(1)}}{2} - \frac{\bar{U} - \bar{V}}{2} \right]. \quad (7)$$

Similarly, the MVUE of the fixed delay δ and mean link delays α and β are the same as in (3). For the sake of completion, the MVUE is also given when α and β are known. It is straightforward to see from (4) that $U_{(1)}$ and $V_{(1)}$ are the complete minimal sufficient statistic for estimating δ and ϕ. The only unbiased functions of $\{U_{(1)}, V_{(1)}\}$ yielding $\boldsymbol{\Phi}_{\mathbf{MVUE}}^{\mathbf{A-K}}$ are

$$\hat{\boldsymbol{\Phi}}_{\mathbf{MVUE}}^{\mathbf{A-K}} = \begin{bmatrix} \hat{\delta}_{\mathrm{MVUE}}^{A-K} \\ \hat{\phi}_{\mathrm{MVUE}}^{A-K} \end{bmatrix}$$
$$= \frac{1}{2} \begin{bmatrix} \left(U_{(1)} - \frac{\alpha}{N}\right) + \left(V_{(1)} - \frac{\beta}{N}\right) \\ \left(U_{(1)} - \frac{\alpha}{N}\right) - \left(V_{(1)} - \frac{\beta}{N}\right) \end{bmatrix}. \quad (8)$$

B. Symmetric Link Delays

In the symmetric case when $\alpha = \beta \triangleq \lambda$, the likelihood function for the clock offset as a function of observations $\{U_k\}_{k=1}^{N}$ and $\{V_k\}_{k=1}^{N}$ is

$$L(\delta, \phi, \lambda) = \lambda^{-2N} \exp\left[-\frac{1}{\lambda} \sum_{k=1}^{N} \{U_k + V_k - 2\delta\}\right]$$
$$\cdot u\left[U_{(1)} - \delta - \phi\right]$$
$$\cdot u\left[V_{(1)} - \delta + \phi\right]. \quad (9)$$

Apparently, for unknown λ, it seems that $\{\sum_{k=1}^{N} U_k, U_{(1)}, \sum_{k=1}^{N} V_k, V_{(1)}\}$ are again the sufficient statistics for the estimation of $\boldsymbol{\Phi}_{\mathbf{MVUE}}^{\mathbf{S-U}} = [\delta \phi \lambda]^T$. But then they have already generated *two* unbiased clock offset estimators, given by (2) and (3). Naturally, this question arises: since the same sufficient statistics have been proved complete, how can they yield two unbiased estimators? Using the factorization theorem, it turns out that $\mathbf{T} = \{\sum_{k=1}^{N}(U_k + V_k), U_{(1)}, V_{(1)}\}$, and not $\{\sum_{k=1}^{N} U_k, U_{(1)}, \sum_{k=1}^{N} V_k, V_{(1)}\}$, is actually the minimal sufficient statistic according to Neymann-Fisher Factorization theorem. Consequently, the clock offset estimator in (3) can not be considered since it is not a function of $\mathbf{T}$.

Now proceeding similarly as before, $r = \sum_{k=1}^{N}(U_k - U_{(1)} + V_k - V_{(1)})$ is Gamma distributed with parameters $(2(N-1), \lambda)$ and

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \pi\left(r, U_{(1)}, V_{(1)}\right)$$
$$\cdot r^{2N-3}$$
$$\times e^{-\frac{N}{\lambda}\left\{\frac{r}{N} + U_{(1)} + V_{(1)}\right\}} dr \, dU_{(1)} \, dV_{(1)} = 0$$
$$\forall \boldsymbol{\Phi}_{\mathbf{MVUE}}^{\mathbf{S-U}}.$$

It follows that $\mathbf{T}$ is also the minimal sufficient statistic and the MVUE is the same as $\hat{\boldsymbol{\Phi}}_{\mathbf{BLUE-OS}}^{\mathbf{S}}$ in (2) expressed as

$$\hat{\boldsymbol{\Phi}}_{\mathbf{MVUE}}^{\mathbf{S-U}} = \begin{bmatrix} \hat{\delta}_{\mathrm{MVUE}}^{S-U} \\ \hat{\phi}_{\mathrm{MVUE}}^{S-U} \\ \hat{\lambda}_{\mathrm{MVUE}}^{S-U} \end{bmatrix}$$
$$= \frac{1}{2(N-1)} \begin{bmatrix} N\left(U_{(1)} + V_{(1)}\right) - (\bar{U} + \bar{V}) \\ (N-1)\left(U_{(1)} - V_{(1)}\right) \\ N\left\{(\bar{U} + \bar{V}) - \left(U_{(1)} + V_{(1)}\right)\right\} \end{bmatrix}. \quad (10)$$

Hence, the MVUE for the clock offset, in the case of symmetric unknown network delays, is expressed as

$$\hat{\phi}_{\mathrm{MVUE}}^{S-U} = \frac{U_{(1)} - V_{(1)}}{2}. \quad (11)$$

Furthermore, the MVUEs for the fixed delay δ and mean link delay λ under the symmetric assumption match the ones in (2). Finally, following a similar procedure, when λ is known, the sufficient statistics are $U_{(1)}$ and $V_{(1)}$ and the MVUE is given by

$$\hat{\boldsymbol{\Phi}}_{\mathbf{MVUE}}^{\mathbf{S-K}} = \begin{bmatrix} \hat{\delta}_{\mathrm{MVUE}}^{S-K} \\ \hat{\phi}_{\mathrm{MVUE}}^{S-K} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} U_{(1)} + V_{(1)} \\ U_{(1)} - V_{(1)} \end{bmatrix}. \quad (12)$$

TABLE I
MVUE FOR THE CLOCK OFFSET

Clock Offset		Delay Mean Known	Delay Mean Unknown
	MVUE	$\left(U_{(1)} - V_{(1)}\right)/2$	$\left(U_{(1)} - V_{(1)}\right)/2$
Symmetric Delays	MSE	$\lambda^2/2N^2$	$\lambda^2/2N^2$
	Remarks	Same as MLE	Same as MLE and BLUE-OS
	MVUE	$\left[\left(U_{(1)} - \frac{\alpha}{N}\right) - \left(V_{(1)} - \frac{\beta}{N}\right)\right]/2$	$\left[N\left(U_{(1)} - V_{(1)}\right) - \left(\overline{U} - \overline{V}\right)\right]/2\left(N-1\right)$
Asymmetric Delays	MSE	$(\alpha^2 + \beta^2)/4N^2$	$(\alpha^2 + \beta^2)/4N(N-1)$
	Remarks	Bias-compensated MLE	Same as BLUE-OS

VI. SIMULATIONS AND EXPLANATORY REMARKS

Summarizing the results derived so far, Table I shows the MVUE for the clock offset for the possible combinations of symmetries/asymmetries in the network delays and knowledge of the mean link delay parameters from (6), (8), (10), and (12).

It is evident from Table I that in practical scenarios where the means of the exponentially distributed delays are unknown, the MVUE is given by (7) or (11) depending on whether the network delays are asymmetric or symmetric. A natural question arises at this stage: which estimator is better when these network delays are *slightly* asymmetric. To answer this question, note that the MVUE is not always the best estimator, it is only the best among unbiased estimators. If some estimator is devised having reduced variance with relatively lesser resultant increase in squared bias, then it can outperform the MVUE in the MSE sense. Hence, for the *asymmetric* unknown mean link delays case, we will compare the MSE of the MLE in (1) with the MVUE in (7) as follows:

$$\text{MSE}\left(\phi_{\text{MVUE}}^{A-U}\right) = \frac{1}{4N(N-1)}(\alpha^2 + \beta^2) \tag{13}$$

$$\text{MSE}(\phi_{\text{MLE}}) = \frac{1}{4N^2}(\alpha^2 + \beta^2) + \frac{1}{4N^2}(\alpha - \beta)^2$$
$$= \frac{1}{2N^2}(\alpha^2 + \beta^2 - \alpha\beta). \tag{14}$$

Notice that though ϕ_{MLE} is biased in the most realistic setting, i.e., asymmetric unknown mean link delays, in accordance with (13) and (14), it outperforms the MVUE under the condition

$$\frac{1}{4N(N-1)}(\alpha^2 + \beta^2) > \frac{1}{2N^2}(\alpha^2 + \beta^2 - \alpha\beta) \tag{15}$$

which can be expressed equivalently as

$$\frac{N}{2} - 1 < \frac{\alpha\beta}{(\alpha - \beta)^2} \triangleq f(\alpha, \beta). \tag{16}$$

The above relations bring into attention a number of remarks. First, (16) provides the number of timing synchronization messages N to be exchanged given α and β, up to which the MLE has lesser MSE than the MVUE for asymmetric link delays. It also suggests that though the MLE is equal to the MVUE *only* in the symmetric link delays case, it attains lesser MSE in the asymmetric case in the region around the point $\alpha = \beta$. As the asymmetry of the link increases, i.e., $|\alpha - \beta|$ tends to drift away

from zero, the MVUE starts outperforming the MLE. The exact point where their performance is the same can be easily derived from (16). The two respective MSEs are drawn in Fig. 2, where N and α are held constant at 15 and 2, respectively, while β is varied across α through the relation $\beta = (\alpha - r/2) : (\alpha + r/2)$. For this plot, the range r is chosen to be 4 and the step size is $r/70$. It shows that the MSE of MLE actually *decreases* when β initially approaches α because the chosen $\alpha = 2$ is a small value and hence the MSE rise due to a slight increase in β is overcome by the fall in the MSE due to the smaller $|\alpha - \beta|$ (for larger values of α, this fall might not occur). It is clear that around the region where $\alpha = \beta$ (illustrated by the solid line in Fig. 2), the MLE outclasses the MVUE and then a further increase in β again results in higher asymmetry thus making the MVUE the better choice. Second, it is evident from (13) and (14) that for a constant N, and increasing α and/or β, the MLE again exhibits better performance than the MVUE, and hence should be preferred over MVUE in networks with large delays. Third, (16) shows that for any $\alpha \neq \beta$, N can be made large enough to surpass the expression on the right hand side. This fact is also clear from Fig. 3, where the same plot is drawn with N ranging from 15 to 20. Notice that although the MSEs of both estimators decrease with N, the two lines representing the intersections of the MSE curves manifest decreasing separation between them. This result corroborates the fact that MVUE overtakes the MLE after a certain number of observations. Fourth, it apparently seems that for a constant N, estimating α and β utilizing (6) and (10) and plugging in (16) might be a good idea for adaptively selecting between the MVUE and MLE as

$$\frac{N}{2} - 1 \underset{\text{MLE}}{\overset{\text{MVUE}}{\gtrless}} \frac{\hat{\alpha}_{\text{MVUE}}^{A-U}\hat{\beta}_{\text{MVUE}}^{A-U}}{\left(\hat{\alpha}_{\text{MVUE}}^{A-U} - \hat{\beta}_{\text{MVUE}}^{A-U}\right)^2} = f\left(\hat{\alpha}_{\text{MVUE}}^{A-U}, \hat{\beta}_{\text{MVUE}}^{A-U}\right).$$

However, since $f(\hat{\alpha}, \hat{\beta})$ processes nonlinearly the estimates, an amplification of estimation errors might occur which affects the quality of the resultant $\hat{f}(\alpha, \beta)$. In other words, even having access to $\hat{\alpha}_{\text{MLE}}$ and $\hat{\beta}_{\text{MLE}}$ might not help to estimate accurately $\hat{f}(\alpha, \beta)$, despite the fact that the MLE is functionally invariant. Nevertheless, such a technique can be used when the asymmetry between the delays is large, since the incorrect choice appears only around the region where the two MSE curves (as in Fig. 2) intersect with each other.

To address briefly the issue pertaining to the robustness of proposed estimators to slight deviations from the exponential
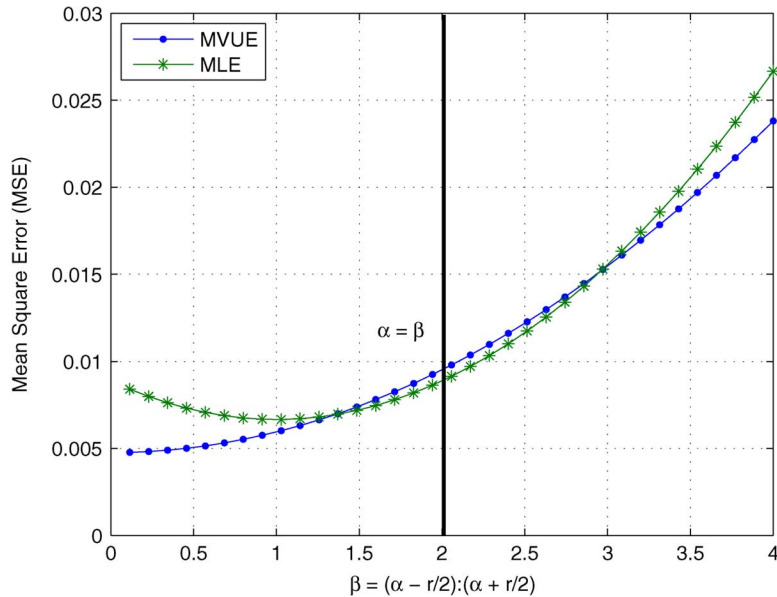
Fig. 2. MSE of the MLE (1) and the MVUE (7) for asymmetric unknown delays with constant $N = 15$.

delay model, we have run computer simulations assuming the same timing message exchange mechanism but with Gamma distributed delays with shape parameter $k = 2$ and scale parameter $\lambda = 1$ (Gamma distribution is one of the most general distributions used for modeling network delays as explained in Section III). There is no closed-form expression for the maximum likelihood estimates of the clock offset and deterministic delay in this case, and the MLE of the clock offset had to be estimated numerically. The results are drawn in Fig. 4 showing that the MVUE derived for exponential case is robust in the presence of asymmetric delays and performs close to the maximum likelihood estimate in Gamma delays. The cost of higher MSE of the former saves a very large number of iterations required for the latter.

However, for distributions quite different than the exponential distribution, the derived MVUE is not robust, as described in detail in [20]. In this case, one can resort to an MMSE particle filtering based approach as described in [21]. An attractive feature of the proposed MMSE particle filtering estimator [21] is the fact that the density of network delays is estimated adaptively in the same time with the unknown clock parameters, and provided that a large number of samples are available, the resulting estimator will be MMSE and very robust to any type of network delay distributions. However, its computational complexity is high in general for WSNs characterized by reduced memory, energy and computational power. Nonetheless, it can be implemented in traditional computer networks and its applicability depends on the complexity-performance tradeoff regime that we want to operate on.

To illustrate the performance of proposed estimators with respect to the existing state-of-the-art estimators, we have performed comparisons with the parametric and non-parametric bootstrap-bias correction (PBC and NBC, respectively) estimators. The procedures for building the bootstrap bias correction estimators are detailed in the Appendix (for interested readers, [22] and [23] are excellent references).

The results of the findings are as follows. In the symmetric exponential model drawn in Fig. 5 for $\lambda = 1$, the MLE exhibits the best performance among those estimators. Notably, in the case that there is no bias, applying a bias correction might result in slight deterioration in MSE performance compared to the MLE. However, in the asymmetric exponential model drawn in Fig. 6 for $\alpha = 1$ and $\beta = 3$, PBC and MVUE present almost the same performance, though PBC is slightly better than the MVUE at the cost of much higher computational complexity.

These findings are very important in the context of WSNs where the energy resources are limited and the number of synchronization packet exchanges is rather small. Even in the traditional centralized or *ad-hoc* networks, it should be noted that for an α fairly close to β, the MLE gives better results no matter by how much magnitude N is increased. In addition, when the topology of the network does not remain constant for longer periods of time as in *ad-hoc* networks, different delay environments are present during different synchronization cycles and choosing between the MVUE and the MLE according to each situation yields a better solution.

Based on the above observations, it should be emphasized that the problem under study provides a very illustrative example about the worth of the MLE in real world scenarios. It is not only relatively easier to derive, but it also performs outstandingly well in comparison to other laboriously obtained optimal (in some sense) estimators. This represents another justification for the widespread usage of MLE in engineering applications.

As a final remark, note that the primary interest of this research was on deriving the estimates for the clock offset but as a byproduct, the estimates of both fixed and variable link delays have also been obtained in (6) and (10), where their BLUE-OS matches again the MVUE. This outcome is also helpful since end-to-end delay measurements are frequently used in analyzing network performance. For example, packet delay statistics are important in examining the performance and reliability of the Internet, but it has no mechanism for providing
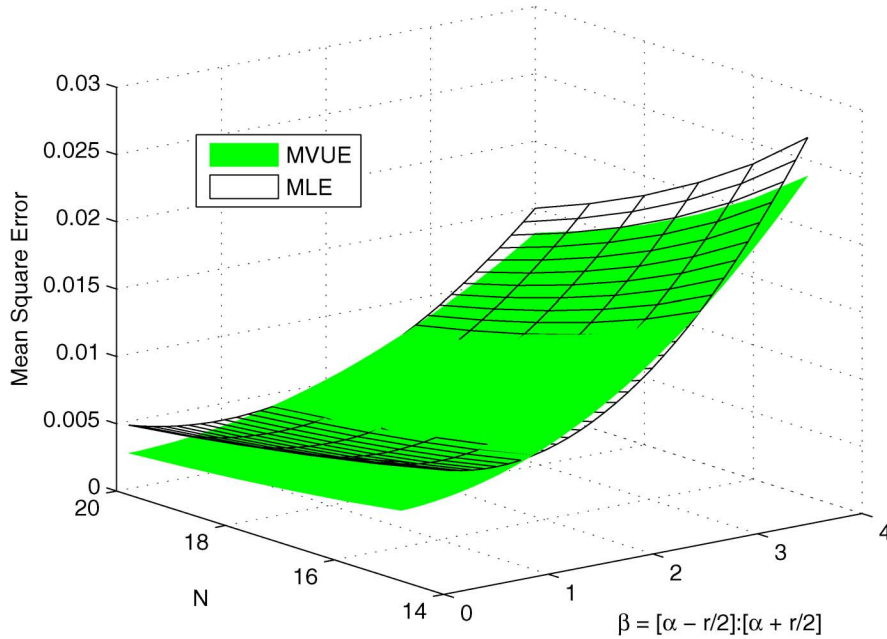
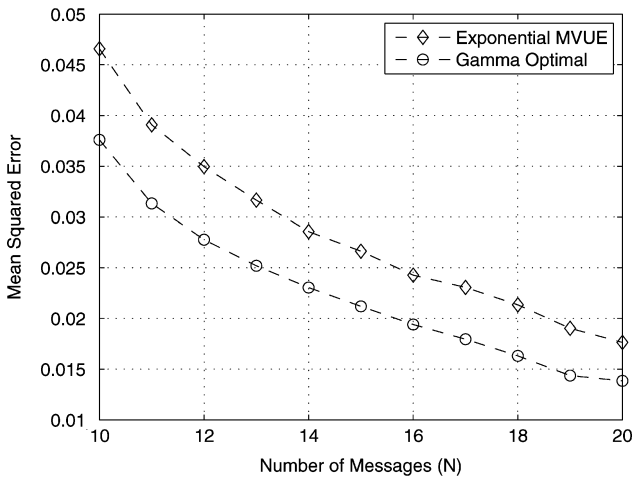Fig. 3.   MSE of the MLE (1) and the MVUE (7) for asymmetric unknown delays with different values of N.

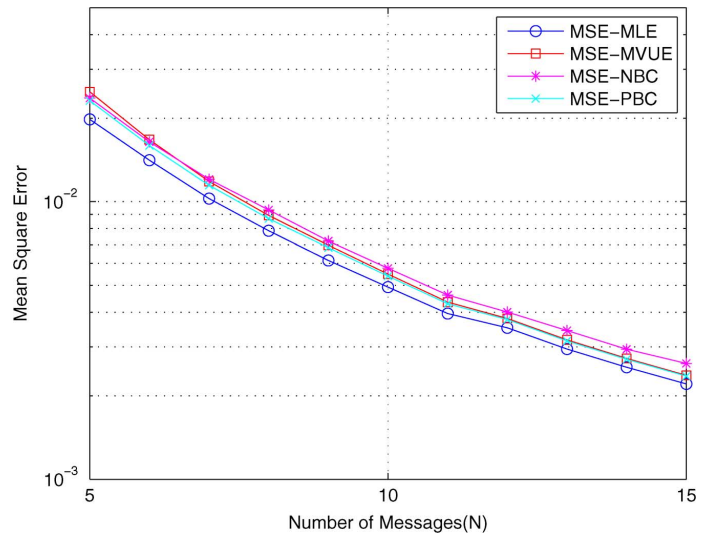Fig. 4.   MSE of the Gamma MLE and the MVUE (7) for unknown delays with different values of N.

Fig. 5.   MSE for Symmetric Exponential Delay Model ($\alpha = \beta = 1$).

feedback on network congestion to end-systems at the IP layer. Moreover, these results are also useful for end-system protocols and applications that behave adaptively based on their control on the observed network performance. Lastly, the estimates of fixed and mean variable delays are also important in other areas such as continuous-media applications, e.g., audio and video applications need to coup with the delay jitter perceived at the receiver for smooth playout of the original stream. For better performance of such applications, determining the correct amount of buffering, and the reconstruction of the original timing plays a significant role.

VII. CONCLUSIONS AND FUTURE WORK

This paper targets the clock synchronization problem in a general two-way timing packets exchange scenario. The best linear unbiased estimate (using order statistics) of the clock offset between two nodes has been derived for both symmetric and asymmetric exponential link delay circumstances. The MVUE is then obtained, and it is shown to coincide with the best linear unbiased estimate using order statistics in each case. However, it matches the MLE only in the symmetric link delay situation. Finally, the maximum likelihood estimator is established to outperform the MVUE in the region around the point of symmetry, thus establishing its authority as the most extensively employed estimation scheme for practical settings.

As a future work, the MMSE estimator, if realizable, is a nice problem to tackle so that a uniformly better solution can be obtained. Another possible direction might be to devise an efficient method for selecting between the above mentioned estimators to enhance the overall performance.
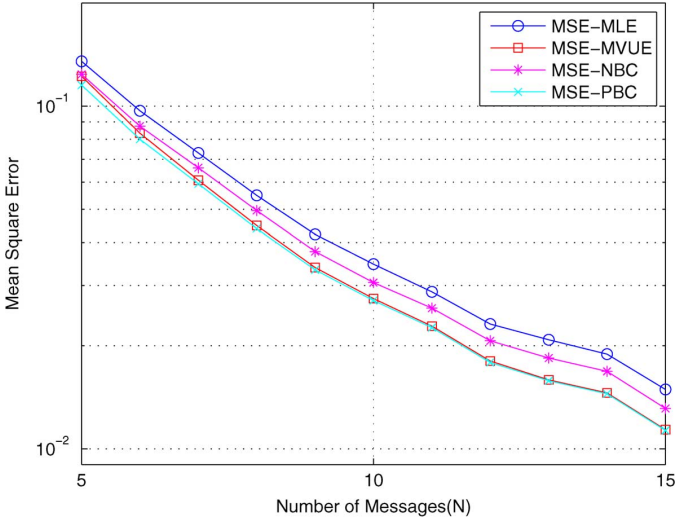
Fig. 6. MSE for Asymmetric Exponential Delay Model ($\alpha = 1$, $\beta = 3$).

APPENDIX

A. Nonparametric Bootstrap

Step 0. Conduct the experiment to obtain the random sample $X = \{X_1, X_2, \ldots, X_n\}$ and calculate the estimate $\hat{\theta}$ from the sample X.

Step 1. Construct the empirical distribution $\hat{F}$, which puts equal mass $1/n$ at each observation $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$.

Step 2. From $\hat{F}$, draw a sample $X^* = \{X_1^*, X_2^*, \ldots, X_n^*\}$, called the bootstrap resample.

Step 3. Approximate the distribution of $\hat{\theta}$ by the distribution of $\hat{\theta}^*$ derived from the bootstrap resample X^*.

B. Parametric Bootstrap

Suppose that one has some partial information about F. For example, F is known to be the exponential distribution but with unknown mean μ. This suggests that we should draw a resample of size n from the exponential distribution with mean $\hat{\mu}$ where $\hat{\mu}$ is estimated from X rather than from a non-parametric estimate $\hat{F}$ of F. We use the exponential distribution in the suggested bias correction through parametric bootstrapping. The parametric bootstrap principle is almost the same as the above nonparametric bootstrap principle, except some steps.

C. Bootstrap Estimate of Bias

Let us suppose that an unknown probability distribution F assumes the data $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ by random sampling, $F \rightarrow x$. We want to estimate a real-valued parameter $\theta = t(F)$. For now we will assume the estimator to be any statistic $\hat{\theta} = s(\mathbf{x})$. The *bias* of $\hat{\theta} = s(\mathbf{x})$ as an estimate of θ is defined to be the difference between the expectation of $\hat{\theta}$ and the value of the parameter θ

$$\text{bias}_F = \text{bias}_F(\hat{\theta}, \theta) = E_F[s(\mathbf{x})] - t(F). \qquad (17)$$

A large bias is usually an undesirable aspect of an estimator's performance. We can use the bootstrap to assess the bias of any estimator $\hat{\theta} = s(\mathbf{x})$. The *bootstrap estimate of bias* is defined to be the estimate $bias_{\hat{F}}$ obtained by substituting $\hat{F}$ for F

$$\text{bias}_{\hat{F}} = E_F[s(\mathbf{x}^*)] - t(\hat{F}). \qquad (18)$$

For most statistics that arise in practice, the ideal bootstrap estimate $bias_{\hat{F}}$ must be approximated by Monte Carlo simulations. We generate independent bootstrap samples $\mathbf{x}^{*1}, \mathbf{x}^{*2}, \ldots, \mathbf{x}^{*B}$, evaluate the bootstrap replications $\hat{\theta}^*(b) = s(\mathbf{x}^{*b})$, and approximate the bootstrap expectation $E_{\hat{F}}[s(\mathbf{x}^*)]$ by the average

$$\hat{\theta}^*(\,\cdot\,) = \sum_{b=1}^{B} \hat{\theta}^*(b)/B = \sum_{b=1}^{B} s(\mathbf{x}^{*b})/B. \qquad (19)$$

The bootstrap estimate of bias based on the B replications $\hat{\text{bias}}_B$, is (2) with $\hat{\theta}^*(\,\cdot\,)$ substituted for $E_{\hat{F}}[s(\mathbf{x}^*)]$

$$\hat{\text{bias}}_B = \hat{\theta}^*(\,\cdot\,) - t(\hat{F}). \qquad (20)$$

D. Bias Correction

The usual reason why we want to estimate the bias of $\hat{\theta}$ is to correct $\hat{\theta}$ so that it becomes less biased. If $\hat{\text{bias}}$ is an estimate of $\text{bias}_F(\hat{\theta}, \theta)$, then the obvious *bias-corrected estimator* is

$$\bar{\theta} = \hat{\theta} - \hat{\text{bias}}. \qquad (21)$$

Taking $\hat{\text{bias}}$ equal to $\hat{\text{bias}}_B = \hat{\theta}^*(\,\cdot\,) - \hat{\theta}$ gives

$$\bar{\theta} = 2\hat{\theta} - \hat{\theta}^*(\,\cdot\,). \qquad (22)$$

REFERENCES

[1] D. R. Jeske, "On maximum likelihood estimation of clock offset," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 53–54, Jan. 2005.

[2] B. Sadler and A. Swami, "Synchronization in sensor networks: An overview," in *Proc. IEEE Military Commun. Conf. (MILCOM 2006)*, Oct. 2006, pp. 1–6.

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

[4] S. Ganeriwal, R. Kumar, and M. B. Srivastava, "Timing synch protocol for sensor networks," in *Proc. 1st Int. Conf. Embedded Netw. Sens. Syst.*, Los Angeles, CA, Nov. 2003.

[5] F. Cristian, "Probabilistic clock synchronization," in *Distributed Computing*. New York: Springer-Verlag, 1989, vol. 3, p. 146158.

[6] K. Arvind, "Probabilistic clock synchronization in distributed systems," *IEEE Trans. Parallel Distrib. Syst.*, vol. 5, no. 5, p. 474487, May 1994.

[7] K. Romer, "Time synchronization in ad hoc networks," in *Proc. ACM Symp. Mobile Ad Hoc Netw. Comput.*, Oct. 2001, p. 173182.

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

[9] 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.

[10] H. S. Abdel-Ghaffar, "Analysis of synchronization algorithm with time-out control over networks with exponentially symmetric delays," *IEEE Trans. Commun.*, vol. 50, pp. 1652–1661, Oct. 2002.

[11] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.

[12] A. L. Garcia, *Probability and Random Processes for Electrical Engineering*, 2nd ed. Reading, MA: Addison-Wesley, 1993.

[13] C. Bovy, H. Mertodimedjo, G. Hooghiemstra, H. Uijterwaal, and P. Mieghem, "Analysis of end-to-end delay measurements in internet," presented at the The Passive and Active Measurements Workshop, Fort Collins, CO, Mar. 2002.

[14] S. Moon, P. Skelley, and D. Towsley, "Estimation and removal of clock skew from network delay measurements," in *Proc. IEEE INFOCOM Conf. Comput. Commun.*, New York, NY, Mar. 1999.

[15] H. Eriksson, "MBONE: The multicast backbone," *Commun. ACM*, vol. 37, no. 8, Aug. 1994.

[16] V. Paxson, "On calibrating measurements of packet transit times," in *Proc. 7th ACM Sigmetrics Conf.*, Jun. 1998, vol. 26, pp. 11–21.

[17] E. H. Lloyd, "Least-squares estimation of location and scale parameters using order statistics," *Biometrika*, vol. 39, no. 1/2, pp. 88–95, Apr. 1952.

[18] N. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, 2nd ed. New York: Wiley, 1994, vol. 1.

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

[20] K.-L. Noh, Q. Chaudhari, E. Serpedin, and B. 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, Apr. 2007.

[21] J. Kim, J. Lee, E. Serpedin, and K. Qaraqe, "A robust estimation scheme for clock phase offsets in wireless sensor networks in the presence of non-Gaussian random delays," *Signal Process.*, vol. 89, no. 6, Jun. 2009, Elsevier.

[22] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*. London, U.K.: Chapman and Hall, 1993.

[23] A. M. Zoubir and D. R. Iskander, *Bootstrap Techniques for Signal Processing*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

Qasim M. Chaudhari received the B.E. degree in electrical engineering from the National University of Sciences and Technology, Pakistan, in 2001. He received the M.S. degree from the University of Southern California, Los Angeles, in 2004, and the Ph.D. degree from Texas A&M University, College Station, in 2008, both in electrical engineering.

He was with Communications Enabling Technologies from 2001 to 2002 in the SoC Tools group. He has worked as an intern in Qualcomm Inc. in HSDPA performance test team in 2005/06. Currently, he is holding an Assistant Professor position in the Department of Electrical Engineering at Iqra University, Pakistan. His research interests include estimation and detection theory in general and synchronization in wireless networks in particular.

Erchin Serpedin received the specialization degree in signal processing and transmission of information from Ecole Superieure D'Electricite (SUPELEC), Paris, France, in 1992, the M.Sc. degree from the Georgia Institute of Technology, Atlanta, in 1992, and the Ph.D. degree in electrical engineering from the University of Virginia, Charlottesville, in January 1999.

In July 1999, he joined Texas A&M University, College Station, where he currently holds the position of Associate Professor. His research interests are in the areas of signal processing and telecommunications. He is also a coauthor of the research monograph, *Synchronization in Wireless Sensor Networks*, (Cambridge, U.K.: Cambridge University Press, August 2009).

Dr. Serpedin received the NSF Career Award in 2001, the Outstanding Faculty Award in 2004, and several other awards. He served as a Technical Co-Chair of the Communications Theory Symposium at Globecom 2006 Conference, and for the VTC Fall 2006: Wireless Access Track Symposium. He served as an Associate Editor for the IEEE COMMUNICATIONS LETTERS, IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE SIGNAL PROCESSING LETTERS, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and EURASIP *Journal on Advances in Signal Processing*. He is currently serving as an Associate Editor for *Signal Processing-Elsevier*, IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON INFORMATION THEORY, and the EURASIP *Journal on Bioinformatics and Systems Biology*.

Khalid Qaraqe (M'97–SM'00) was born in Bethlehem. He received the B.S. degree in electrical engineering from the University of Technology, Baghdad, in 1986, with honors. He received the M.S. degree in electrical engineering from the University of Jordan, Jordan, in 1989, and the Ph.D. degree in electrical engineering from Texas A&M University, College Station, in 1997.

From 1989 to 2004, he held a variety of positions in many companies and he has more than 12 years of experience in the telecommunication industry. He has worked for Qualcomm, Enad Design Systems, Cadence Design Systems/Tality Corporation, STC, SBC, and Ericsson. He has worked on numerous GSM, CDMA, and WCDMA projects and has experience in product development, design, deployments, testing, and integration. He joined the Department of Electrical Engineering, Texas A&M University at Qatar, in July 2004, where he is now an Associate Professor. His research interests include communication theory and its application to design and performance, analysis of cellular systems, and indoor communication systems. Of particular interest is the development of WCDMA and broadband wireless communications and diversity techniques.