

ESTIMATION OF CLOCK PARAMETERS AND PERFORMANCE
BENCHMARKS FOR SYNCHRONIZATION IN WIRELESS SENSOR
NETWORKS

A Dissertation

by

QASIM MAHMOOD CHAUDHARI

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

August 2008

Major Subject: Electrical Engineering

ESTIMATION OF CLOCK PARAMETERS AND PERFORMANCE
BENCHMARKS FOR SYNCHRONIZATION IN WIRELESS SENSOR
NETWORKS

A Dissertation

by

QASIM MAHMOOD CHAUDHARI

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	Erchin Serpedin Khalid Qaraqe
Committee Members,	Jean-Francois Chamberland Narasimha Reddy Jennifer Welch
Head of Department,	Costas Georgiades

August 2008

Major Subject: Electrical Engineering

ABSTRACT

Estimation of Clock Parameters and Performance Benchmarks for Synchronization
in Wireless Sensor Networks. (August 2008)

Qasim Mahmood Chaudhari,

B.E., National University of Sciences and Technology, Rawalpindi, Pakistan

M.S., University of Southern California, Los Angeles

Co-Chairs of Advisory Committee: Dr. Erchin Serpedin
Dr. Khalid Qaraqe

Recent years have seen a tremendous growth in the development of small sensing devices capable of data processing and wireless communication through their embedded processors and radios. Wireless Sensor Networks (WSNs) are ad hoc networks consisting of such devices gaining importance due to their emerging applications. For a meaningful processing of the information sensed by WSN nodes, the clocks of these individual nodes need to be matched through some well defined procedures. This dissertation focuses on deriving efficient estimators for the clock parameters of the network nodes for synchronization with the reference node and the estimators variance thresholds are obtained to lower bound the maximum achievable performance.

For any general time synchronization protocol involving a two way message exchange mechanism, the BLUE-OS and the MVUE of the clock offset between them is derived assuming both symmetric and asymmetric exponential network delays. Next, with the inclusion of clock skew in the model, the joint MLE of clock offset and skew under both the Gaussian and the exponential delay model and the corresponding algorithms for finding these estimates are presented. Also, for applications where even clock skew correction cannot maintain long-term clock synchronization, a closed-form

expression for the joint MLE for a quadratic model is obtained.

Although the derived MLEs are not computationally very complex, two computationally efficient algorithms have been proposed to estimate the clock offset and skew regardless of the distribution of the delays. Afterwards, extending the idea of having inactive nodes in a WSN overhear the two-way timing message communication between two active (master and slave) nodes, the MLE, the BLUE-OS, the MVUE and the MMSE estimators for the clock offsets of the inactive nodes located within the communication range of the active nodes are derived, hence synchronizing with the reference node at a reduced cost.

Finally, focusing on the the one-way timing exchange mechanism, the joint MLE for clock phase offset and skew under exponential noise model and the Gibbs Sampler for a receiver-receiver protocol is formulated and found via a direct algorithm. Lower and upper bounds for the MSE of JMLE and Gibbs Sampler are introduced in terms of the MSEs of the MVUE and the conventional BLUE, respectively.

To Sikandar Shaheen

ACKNOWLEDGMENTS

I feel deeply indebted to my parents, my wife and family for their constant encouragement and prayers that helped me accomplish this goal. I would also like to thank my advisor, Dr. Erchin Serpedin, who not only supervised my research with exceptional commitment, but also supported me in every possible manner.

In addition, I want to express my gratitude for my co-advisor, Dr. Khalid Qaraqe, and my committee members, Dr. Jean-Francois Chamberland, Dr. Jennifer Welch, Dr. Narasimha Reddy and the Department Head Dr. Costas Georghiadis for their academic guidance. I must also acknowledge my colleagues Yi Zhou, Kyoung-Lae Noh, Huseyin Peksen, Jae-Han Lee, Sangwoo Park and Ilkay Sari, with whom I had a great time in College Station.

Finally, I like to dedicate this dissertation to my childhood friend Sikandar Shaheen, whose demise at such a young age of 25 taught me the truth about life.

TABLE OF CONTENTS

CHAPTER		Page
I	INTRODUCTION	1
	A. Time Synchronization in Distributed Systems	1
	B. Time Synchronization in Wireless Sensor Networks	3
	C. Sources of Error in Time Synchronization	6
	D. Contributions of This Research	8
II	BLUE-OS AND MVUE FOR CLOCK OFFSET	12
	A. The System Architecture	13
	B. Best Linear Unbiased Estimation Using Order Statistics (BLUE-OS)	15
	1. Symmetric Link Delays	18
	2. Asymmetric Link Delays	20
	C. Minimum Variance Unbiased Estimation (MVUE)	23
	1. Asymmetric Link Delays	24
	2. Symmetric Link Delays	29
	D. Explanatory Remarks	31
	E. Summary	37
III	CLOCK OFFSET AND SKEW ESTIMATION*	39
	A. Gaussian Delay Model	40
	1. Maximum Likelihood Clock Offset Estimation	41
	2. Cramer-Rao Lower Bound for Clock Offset	41
	3. Joint Maximum Likelihood Estimation of Clock Offset and Skew	42
	4. Cramer-Rao Lower Bound for Clock Offset and Skew	47
	B. Exponential Delay Model	49
	1. Cramer-Rao Lower Bound for Clock Offset	49
	2. Joint Maximum Likelihood Estimation of Clock Offset and Skew	53
	a. Case I: τ known, ϕ known	54
	b. Case II: τ unknown, ϕ known	56
	c. Case III: τ known, ϕ unknown	64
	d. Case IV: τ unknown, ϕ unknown	68

CHAPTER	Page
	74
IV	75
	76
	77
	78
	81
	83
	85
	88
	90
	91
V	93
	94
	98
	111
	113
	118
	119
	121
VI	122
	123
	125
	129
	135
	137
	137
	140
	143
	144

CHAPTER	Page
VII	CLOCK DRIFT ESTIMATION FOR ACHIEVING LONG-TERM SYNCHRONIZATION* 146
	A. Problem Formulation 147
	B. The Estimation Procedure 148
	C. Summary 153
VIII	JOINT SYNCHRONIZATION OF CLOCK OFFSET AND SKEW IN A RECEIVER-RECEIVER PROTOCOL* 155
	A. Modeling Assumptions 155
	B. JML Estimation of the Offset and Skew 156
	C. Application of Gibbs Sampler 158
	D. Performance Bounds and Simulations 160
	E. Summary 164
IX	CONCLUSIONS AND FUTURE WORK 165
	REFERENCES 167
	VITA 173

LIST OF TABLES

TABLE		Page
I	The MVUE for the clock offset for known mean link delay	32
II	The MVUE for the clock offset for unknown mean link delay	32
III	Computational complexity of each algorithm	92
IV	Intersections of surfaces (5.7) - (5.9)	102
V	Boundary evaluation of region R_a	105
VI	Boundary evaluation of region R_b	106
VII	Boundary evaluation of region R_c	107
VIII	Boundary evaluation of region R_d	108
IX	Boundary evaluation of region R_e	109
X	Boundary evaluation of region R_f	110
XI	Boundary evaluation of region R_g	111

LIST OF FIGURES

FIGURE	Page
1	A sender-receiver timing message exchange paradigm. 14
2	Mean square error of the MLE (2.1) and the MVUE (2.9) for asymmetric unknown delays with constant $N = 15$ 35
3	Mean square error of the MLE (2.1) and the MVUE (2.9) for asymmetric unknown delays with different values of N 36
4	CRLB and variance of the MLE of clock offset for the Gaussian delay model ($\sigma = 1$). 42
5	Two-way timing message exchange model having clock offset and skew. 43
6	MSE of clock offset estimate $\hat{\phi}$ as a function of variance of delay. . . 47
7	CRLB and variance of the MLE of clock offset for the exponential delay model ($\alpha = 1$). 51
8	Variances of both MLEs of clock offset for exponential and Gaussian delays ($\alpha = 1$ and $\sigma = 0.5$). 52
9	Nonzero likelihood region shown as the solid line. 55
10	Support region of the likelihood function shown as the shaded area. Also shown to the left in this figure is the sign of the term $2NT_{3,j} - \sum_{k=1}^N (T_{3,k} - T_{2,k})$ for each $j = 1, \dots, N$ 57
11	Zoomed in version of the support region of the likelihood function. . . 60
12	τ as a function of ϕ and ω 65
13	ω as a function of ϕ for constant τ 66
14	τ as a function of ω only. 69
15	τ as a function of ϕ only. 70

FIGURE	Page
16	Regions of the order statistics $\{\delta_{(i)}\}_{i=1}^2$ 80
17	Variance of the MLE of the Gaussian delay model (GMLE) and the Gaussian MLLE (GMLLE) for Gaussian random delays ($\sigma = 1$). 84
18	Variance of the GLME and the exponential MLLE (EMLLE) for exponential random delays ($\alpha = 1$). 84
19	Variance of the GLME and the MLLEs for Gamma random delays ($\gamma = 2$). 85
20	Variance of the joint ML clock offset estimate and the proposed estimator for Gaussian random delays ($\sigma = 0.5$). 86
21	The estimated fit with the original curve. 87
22	Comparison of the proposed algorithm with case IV. 89
23	Comparison of the proposed algorithm with case IV for Gamma distributed random delays. 90
24	A WSN with two active nodes m and p exchanging timing cells with silent nodes like q , located within the common broadcast region of the active nodes m and p 95
25	A two-way timing cell exchange mechanism between nodes m and p with node q overhearing them. 96
26	Support region of the reduced likelihood function $L'(\tau, \phi_p, \phi_q)$ 99
27	An aerial 2-D view of the support region 101
28	A breakdown of the support region in 7 parts 102
29	Simulations illustrating the MSE of $\hat{\phi}_q$ and $\hat{\phi}_p$ on a logarithmic scale for $\lambda = 1$ 120
30	\mathbf{S} and the solution s_1 157
31	MSE for $\hat{\theta}_{x,BLUE}$, $\hat{\theta}_{x,JML}$, $\hat{\theta}_{x,GIBBS}$, and $\hat{\theta}_{x,MVUE}$ 162
32	MSE for $\hat{\beta}_{x,BLUE}$, $\hat{\beta}_{x,JML}$, $\hat{\beta}_{x,GIBBS}$ and $\hat{\beta}_{x,MVUE}$ 163

CHAPTER I

INTRODUCTION

A. Time Synchronization in Distributed Systems

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 complex problems of computer engineering. Whether it is the disciplining of computer clocks with the devices synchronized to a GPS satellite or a Network Time Protocol (NTP) time 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 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

¹The journal model is *IEEE Transactions on Automatic Control*.

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. During the last two decades, many clock synchronization protocols have been proposed such as [1], [2], [3], etc. The Network Time Protocol (NTP) [1] 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 architecture is widely accepted in designing time synchronization algorithms and consists of the same two-way timing message exchange mechanism targeted in most of this dissertation.

A protocol based on the remote clock reading method was put forward by [2], which handles unbounded message delays between processes. In [3], 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 [4] represents one of the pioneering contributions in this area. The protocol is based on generating timestamps to record the time at which an event of interest occurred. 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.

B. Time Synchronization in Wireless Sensor Networks

Recent technological advances have made it possible to design miniature devices (sensors) capable of performing onboard sensing, computing and communication tasks. A Wireless Sensor Network (WSN) consists of a large number of such tiny devices, called nodes, that are connected in an ad hoc manner without assuming any centralized infrastructure [5]. Since the WSN nodes are deployed in an ad hoc fashion and mostly left without any maintenance and battery replacement for their lifetimes, they are usually cheap and hence unreliable. Therefore, all the design aspects of a sensor network concentrate on minimizing energy utilization [6]-[8]. With as peculiar characteristics as limited energy sources, high density of node deployment and cheap and unreliable sensor nodes, sensor networks are designed to perform complex tasks such as:

- Environment: Flood detection, forest fire monitoring and ecological and biological habitats.
- Military: Presence of hazardous materials, monitoring equipment and ammunition.
- Science: Deep sea exploration, study of cosmic radiation.
- Civil: Surveillance for security in shopping malls and banks, traffic monitoring.

For successfully performing most of their applications, time synchronization in wireless sensor networks is very important due to a number of reasons:

- Sensor nodes need to coordinate their operations and collaborate to achieve a complex sensing task. Data fusion is an example of such coordination in which

data collected at different nodes are aggregated into a meaningful result. For example, in a vehicle tracking application, sensor nodes report the location and time at which they sense the vehicle to a sink node that in turn combines this information to estimate the location and velocity of the vehicle. Clearly, if the sensor nodes are not synchronized, the estimates will be inaccurate. Similarly, the integration of voice, video or environmental data from different sensors and its processing in a meaningful way requires the node to be synchronized in time.

- Scheduling algorithms such as TDMA can be used to share the transmission medium in the time domain to eliminate transmission collisions and conserve energy.
- Time synchronization enables all the nodes in the network to assume efficient duty cycling operation, i.e., coordinated sleep and wake up modes. For example, sensor may go into power saving mode by turning off their sensors and/or transceivers. When running in these modes, they should sleep and wake up at coordinated times, such that radio receiver of a node is not turned off when it has to participate in a pre-assigned activity. This requires precise timing between sensor nodes.
- Moreover, many localization, security and tracking protocols also demand the nodes to timestamp their messages and sensing events.

There are a few methods through which the accuracy of the nodes' clocks can be improved, e.g., using GPS to synchronize the hardware clocks to a global reference, using precise clock boards for the nodes, etc. But these solutions prove to be fairly expensive or inappropriate when the nodes have to be low-cost and energy efficient. In addition, the sensor nodes may be left unattended for a long period of time, e.g., on the

ocean floor or in deep space. Also, the conventional network synchronization protocols cannot be employed due to the WSN constraints mentioned above [9]. Hence, there is a need for time synchronization protocols specifically designed to the characteristics of WSNs to make them operate under a common time scale.

Time synchronization in WSNs requires designing a protocol in which the nodes exchange messages with each other to adjust their clocks to a common reference. At the same time, it is highly desirable to extract information about their relative frequency, called clock skew, from the same set of message exchange, because imperfections in quartz crystals and environmental conditions cause different nodes to run at different frequencies. Clock skew adjustment guarantees not only a more accurately synchronized network, but also helps in maintaining this synchronization for a longer period. Hence, it significantly reduces the resynchronization period, i.e., the time interval after which the clock difference among the nodes exceeds some set limits and the network has to resynchronize itself, resulting in tremendous reduction in communication overheads and corresponding energy savings for the whole network.

To deal efficiently with the specific requirements associated with the long-term operation of WSNs, quite a few synchronization protocols have been designed in the past few years. Reference Broadcast Synchronization (RBS) [10] is a pioneering work based on the post-facto receiver-receiver synchronization. In RBS, a reference broadcast message is sent by a node to two or more neighboring nodes which record their own local clocks at the reception of broadcasted message. After collecting a few readings, the nodes exchange their observations and a linear regression approach is used to estimate their relative clock offset and skew. Timing Synch Protocol for Sensor Networks (TPSN) [11] is a conventional sender-receiver protocol which assumes two operational stages: the level discovery phase followed by the synchronization phase. During the level discovery phase, WSN is organized in the form of a spanning tree,

and the global synchronization is achieved by enabling each node to get synchronized with its parent (the node located in the adjacent upper level) by means of a message exchange mechanism through adjusting *only* its clock offset. Timing Synchronization protocol for High Latency acoustic networks (TSHL) [12] combines both of these approaches in two stages. The first stage is similar to RBS while the second stage is similar to TPSN, and it is particularly suitable for networks involving high message delays, e.g., underwater acoustic networks. Flooding Time Synchronization Protocol (FTSP) [13] also combines the two approaches in the sense that the beacon node sends its timestamps within the reference broadcast messages.

All of the above mentioned protocols have their own benefits and limitations. Choosing a protocol which corrects only the clock offset (such as TPSN [11]) results in more utilization of power since synchronization has to be done frequently at regular intervals to prevent the clock skew drift the two clocks too far apart. For example, re-synchronization must be performed after every few minutes in TPSN for applications using Berkeley motes [14]. On the other hand, an assumption of simultaneous reception of reference broadcasts is necessary in protocols which correct both the clock offset and skew (such as RBS [10] and FTSP [13]), which is not only a simplification of the correct model but also not applicable in some cases, e.g., in underwater acoustic sensor networks [12].

C. Sources of Error in Time Synchronization

A description of sources of variability was first described by [15], and extended by [16] incorporating physical layer jitter. A brief overview of the sources of non-determinism and hence errors were presented in [12] as below.

1. Send Time: The delay in the packet traversal from the message assembly at the

application layer all the way down to MAC layer.

2. Access Time: The waiting time for accessing the channel after reaching the MAC layer. This is the most significant factor and highly variable according to the specific MAC protocol. The access time is nondeterministic and varies up to hundreds of milliseconds.
3. Interrupt Handling Time: The delay between the radio chip raising and the microcontroller responding to an interrupt. It can be an issue if interrupts are disabled on the microcontroller.
4. Transition and Reception Time: The delay in sending or receiving the entire length of the packet over the channel. Largely deterministic, a function of bandwidth and packet size.
5. Propagation Time: The delay, for a particular symbol of the message, in traversing all the way to the receiver. The propagation time can be deterministic if the speed of propagation is assumed constant, and endpoint location is known.
6. Encoding and Decoding Time: The time taken by the radio chip to encode/decode and transform a part of the message to/from electromagnetic waves. This time is deterministic and is in the order of hundred microseconds [3].
7. Byte Alignment Time: The delay because of the different byte alignment at the receiver. This time is deterministic and can be computed on the receiver side from the bit offset and the speed of the radio wave.
8. Receive Time: Time for the incoming message to traverse up till the receiver application. It is highly variable and varies for each (stack,OS) pair.

Existing time synchronization schemes focus on eliminating or accounting for these sources of error. Schemes typically differ due to differing assumptions, in which sources of variation are dominant in different domains, and due to different approaches to eliminate the sources of error.

D. Contributions of This Research

In 2002, [17] presented a detailed analysis of clock offset estimation for a general sender-receiver protocol assuming a symmetric exponential delay model. It was implicitly argued that for a known fixed delay τ and exponential delay parameter λ , the MLE of clock offset does not exist because the likelihood function does not possess a unique maximum with respect to the clock offset. However, in 2005, it was proved by [18] that for τ unknown, irrespective of λ being known or unknown, the MLE of the clock offset does exist and coincides with a previously proposed estimator in [19] based on experimental data. The current dissertation substantially extends this research on time synchronization problem as follows.

In Chapter II, the BLUE-OS of the clock offset between two nodes for a sender-receiver timing exchange paradigm are derived assuming both symmetric and asymmetric exponential network delays. The Rao-Blackwell-Lehmann-Scheffé theorem is then exploited to obtain the MVUE for the clock offset and 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 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.

Chapter III focuses on analyzing and deriving the MLEs and the corresponding CRLBs for the conventional clock offset model in a sender-receiver timing exchange assuming Gaussian model for the noise. Next, the joint MLE and corresponding CRLB using a more realistic linear clock offset and skew model assuming Gaussian random delays are also obtained. The MLEs for the clock offset only case, and both the clock offset and skew case, under exponential delay assumption are then derived and the corresponding algorithms for finding these estimates are also presented in detail.

Although the MLEs are robust and have a computational complexity still implementable, simpler algorithms even with the sacrifice of some performance grade, are more suited to low power constraints of WSNs. Therefore, Chapter IV discusses two simplified schemes to estimate both the clock offset and skew requiring negligible computations. The first scheme utilizes the first and the last sample of the observations and the estimators are derived under both the Gaussian and exponential delay models, while the second scheme fits a line between two points at minimum distance apart regardless of the actual delay distribution involved. The simulation results for a comparison of performance with the MLE are also presented.

Extending the idea of having inactive nodes in a WSN overhear the two-way timing message communication between two active (master and slave) nodes, Chapter V derives the MLE for the clock offsets of the inactive nodes located within the communication range of the active nodes by assuming an exponential link delay modeling, hence synchronizing with the reference node at essentially zero cost. A vital implication of this work is that the performance of the sender-receiver protocols, whose main disadvantage has always been categorized as the high communication overhead in WSN scenarios due to their point-to-point rather than the broadcast nature, can be compared with that of receiver-receiver protocols on equal grounds. In addition,

the CRLBs for both the active and inactive nodes are also obtained as a performance benchmark.

Chapter VI advances the results from Chapter V in two domains: First, the BLUE-OS is derived by applying general least-squares theory to an ordered sample. The MVUE is also obtained by the application of Rao-Blackwell-Lehmann-Scheffé theorem, which is shown to coincide with the BLUE-OS. In addition, since the MSE is usually selected as the performance criterion in estimation theory, which can be further decreased by adding slight bias to an estimator at a cost of reduced variance, an MMSE estimator with expected loss independent of the clock offset and fixed delay is also derived outperforming the MVUE. Second, the results presented there are generalized by addressing the problem for both symmetric and asymmetric exponential delays, since the practical message exchange involves asymmetry in general for each direction.

Since many applications require long-term synchronization among the nodes, and experiments have shown that a quadratic model of clock variations can better capture the dynamics of the actual clock model involved, the MLE for all the clock parameters in a two-way timing exchange model with exponential delays are derived in Chapter VII.

Focusing on a receiver-receiver protocol, Chapter VIII then obtains the joint MLE for the clock offset and skew under exponential noise model. The Gibbs Sampler is also proposed for joint clock offset and skew estimation and shown to provide superior performance relative to JML-estimator. Lower and upper bounds for the MSEs of the JML-estimator and Gibbs Sampler are introduced in terms of the MSE of the MVUE and the conventional BLUE, respectively.

Finally, Chapter IX summarizes the results of the dissertation with concluding remarks and also formulates some possible future research directions. It should be

noted that the results presented in this dissertation are applicable to a wide range of time transfer problems either directly or through some minor extensions.

CHAPTER II

BLUE-OS AND MVUE FOR CLOCK OFFSET

Assuming both symmetric and asymmetric exponentially distributed network link delays, this chapter focuses on finding the BLUE-OS and the MVUE for the clock offset between two nodes and evaluates their performance in terms of the MSE, which is chosen as the performance criterion throughout this dissertation. The timing exchange mechanism between the two nodes is the same classical two-way message exchange mechanism adopted in protocols such as NTP [1], TPSN [11], etc. The main contributions of this chapter are as follows.

1. A relatively unnoticed estimation scheme in engineering literature, the BLUE-OS, is investigated in the context 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, BLUE-OS 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, it outperforms the MVUE in terms

of the achievable MSE in the region around the point of link symmetry.

A. The System Architecture

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 A and B are shown in Fig. 1. The messages $T_{1,k}$ and $T_{4,k}$ represent the times measured by the local clock of node A , while the messages $T_{2,k}$ and $T_{3,k}$ represent the times measured by the local clock of node B (which is also the reference). The synchronization procedure starts at time $T_{1,1}$ and at each successive message exchange round k , node A sends a synchronization packet containing the timestamp $T_{1,k}$ to node B which records its reception time as $T_{2,k}$. At $T_{3,k}$, node B sends an acknowledgement packet back to node A containing the timestamps $T_{2,k}$ and $T_{3,k}$, which is delivered and timestamped at time $T_{4,k}$ in accordance with node A 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:

$$\begin{aligned} T_{2,k} &= T_{1,k} + \tau + \phi + X_k, \\ T_{4,k} &= T_{3,k} + \tau - \phi + Y_k. \end{aligned}$$

For simplification, the above equations will be rewritten as

$$\begin{aligned} U_k &= \tau + \phi + X_k, \\ V_k &= \tau - \phi + Y_k, \end{aligned}$$

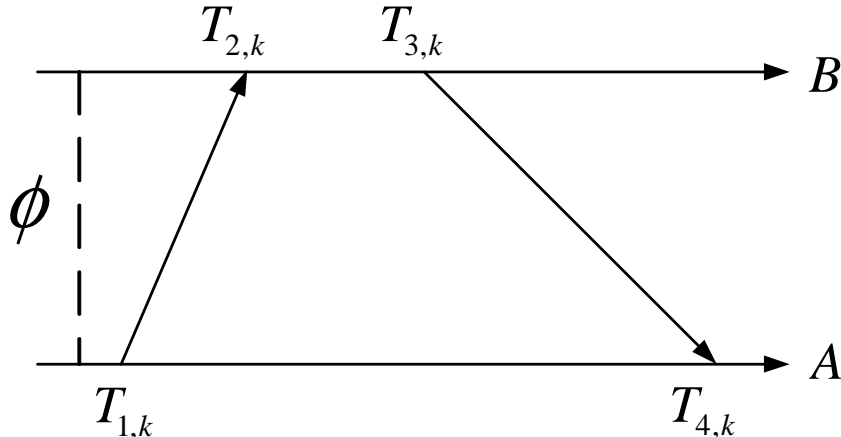


Fig. 1. A sender-receiver timing message exchange paradigm.

where $U_k \triangleq T_{2,k} - T_{1,k}$ and $V_k \triangleq T_{4,k} - T_{3,k}$. The quantity τ symbolizes the fixed portions of the delays assumed to be symmetric for each direction, X_k and Y_k denote the variable portions of delays and assume exponential distributions with means α and β , respectively, and ϕ stands for the clock offset of reference node B with respect to node A .

Network delay modeling has always been an active research topic for the last decades. Out of the proposed probability density function models to capture the distribution of the network delays, the Weibull, exponential, Gamma, and log-normal distributions [20] - [22] have received the greatest attention. There are various reasons behind choosing the exponential distribution for the purpose of this study. Reference [23] collected several traces of delay measurements on the Internet and MBone [24] 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 [17]. Moreover, [17] 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 [19]. Reference [18] 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 for experimental observations.

In [18], it was argued that for an unknown τ , irrespective of the symmetric exponential distribution mean $\alpha = \beta \triangleq \lambda$ being known or unknown, the MLE of the vector parameter $\Phi_{\text{MLE}}^{\text{S}} = [\tau \ \phi \ \lambda]$ is given by

$$\Phi_{\text{MLE}}^{\text{S}} = \begin{bmatrix} \hat{\tau}_{MLE}^{\text{S}} \\ \hat{\phi}_{MLE}^{\text{S}} \\ \hat{\lambda}_{MLE}^{\text{S}} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} U_{(1)} + V_{(1)} \\ U_{(1)} - V_{(1)} \\ \bar{U} + \bar{V} - (U_{(1)} + V_{(1)}) \end{bmatrix}, \quad (2.1)$$

where the subscript S represents the symmetric delay case, $U_{(1)}$ and $V_{(1)}$ denote the minimum order statistics 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 MLE of $\{\tau, \phi\}$ remains the same.

Next, the BLUE-OS and MVUE are derived for both asymmetric and symmetric cases, assuming known and unknown exponential delay means.

B. Best Linear Unbiased Estimation Using Order Statistics (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 modeling framework adopted in this chapter, the application of BLUE is not of much significance. However, for a general location-scale distribution, [25] suggested a new 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

$$\begin{aligned} U'_k &\triangleq \frac{1}{\alpha} (U_k - \tau - \phi), \\ V'_k &\triangleq \frac{1}{\beta} (V_k - \tau + \phi), \end{aligned}$$

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:

$$\begin{aligned} E[U_{(k)}] &= \tau + \phi + \alpha E[U'_{(k)}] \quad , \quad E[V_{(k)}] = \tau - \phi + \beta E[V'_{(k)}] \quad , \\ \text{var}[U_{(k)}] &= \alpha^2 \text{var}[U'_{(k)}] \quad , \quad \text{var}[V_{(k)}] = \beta^2 \text{var}[V'_{(k)}] \quad , \\ \text{cov}[U_{(k)}U_{(j)}] &= \alpha^2 \text{cov}[U'_{(k)}U'_{(j)}] \quad , \quad \text{cov}[V_{(k)}V_{(j)}] = \beta^2 \text{cov}[V'_{(k)}V'_{(j)}] \quad . \end{aligned}$$

Now using standard results from [26], the statistics of the ordered samples can be

expressed as

$$\begin{aligned} E[U'_{(k)}] &= E[V'_{(k)}] = \sum_{i=1}^k \frac{1}{(N-i+1)}, \\ \text{var}[U'_{(k)}] &= \text{var}[V'_{(k)}] = \sum_{i=1}^k \frac{1}{(N-i+1)^2}, \\ \text{cov}[U'_{(k)}U'_{(j)}] &= \text{cov}[V'_{(k)}V'_{(j)}] = \sum_{i=1}^k \frac{1}{(N-i+1)^2}. \end{aligned}$$

As a result, the $N \times N$ symmetric positive-definite covariance matrix \mathbf{C} for both $U'_{(k)}$ and $V'_{(k)}$ takes the form:

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

A simple exercise utilizing Gauss-Jordan elimination yields the following closed-form expression for the inverse of the covariance matrix:

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

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

1. Symmetric Link Delays

The symmetric network delay assumption holds true 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 $\Phi_{\text{BLUE-OS}}^{\text{S}} \triangleq [\tau \ \phi \ \lambda]^T$, which is a linear function of an ordered set of observations $\{U_{(k)}\}_{k=1}^N$ and $\{V_{(k)}\}_{k=1}^N$. Let $\mathbf{z} \triangleq [U_{(1)} \ U_{(2)} \ \cdots \ U_{(N)} \ V_{(1)} \ V_{(2)} \ \cdots \ V_{(N)}]^T$. Then, it is straightforward to notice that

$$E[\mathbf{z}] = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 & -1 & -1 & \cdots & -1 \\ \frac{1}{N} & \frac{1}{N} + \frac{1}{N-1} & \cdots & \sum_{k=1}^N \frac{1}{(N-k+1)} & \frac{1}{N} & \frac{1}{N} + \frac{1}{N-1} & \cdots & \sum_{k=1}^N \frac{1}{(N-k+1)} \end{bmatrix}^T \times \begin{bmatrix} \tau \\ \phi \\ \lambda \end{bmatrix} = \mathbf{Q} \Phi_{\text{BLUE-OS}}^{\text{S}},$$

where \mathbf{z} is the $2N \times 1$ ordered data vector, \mathbf{Q} is a known matrix of dimension $2N \times 3$ and $\Phi_{\text{BLUE-OS}}^{\text{S}}$ is the 3×1 vector of unknown parameters. The above linear relationship lends the problem to be solved by the Gauss-Markov theorem as follows:

$$\hat{\Phi}_{\text{BLUE-OS}}^{\text{S}} = (\mathbf{Q}^T \mathbf{C}_{\mathbf{z}}^{-1} \mathbf{Q})^{-1} \mathbf{Q}^T \mathbf{C}_{\mathbf{z}}^{-1} \mathbf{z}.$$

Since $\{U_{(k)}\}_{k=1}^N$ and $\{V_{(k)}\}_{k=1}^N$ are independent data sets, $\mathbf{C}_{\mathbf{z}}$ is now given by

$$\mathbf{C}_{\mathbf{z}} = \lambda^2 \begin{bmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{bmatrix},$$

and its inverse can be expressed as

$$\mathbf{C}_z^{-1} = \frac{1}{\lambda^2} \begin{bmatrix} \mathbf{C}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^{-1} \end{bmatrix}.$$

It follows that

$$\mathbf{Q}^T \mathbf{C}_z^{-1} \mathbf{Q} = \frac{1}{\lambda^2} \begin{bmatrix} 2N^2 & 0 & 2N \\ 0 & 2N^2 & 0 \\ 2N & 0 & 2N \end{bmatrix},$$

and its inverse is

$$(\mathbf{Q}^T \mathbf{C}_z^{-1} \mathbf{Q})^{-1} = \frac{\lambda^2}{2N(N-1)} \begin{bmatrix} 1 & 0 & -1 \\ 0 & \frac{N-1}{N} & 0 \\ -1 & 0 & N \end{bmatrix}. \quad (2.2)$$

This yields the multiplicative factor of \mathbf{z} as

$$(\mathbf{Q}^T \mathbf{C}_z^{-1} \mathbf{Q})^{-1} \mathbf{Q}^T \mathbf{C}_z^{-1} = \frac{1}{2N(N-1)} \begin{bmatrix} N^2 - 1 & -1 & \cdots & -1 & N^2 - 1 & -1 & \cdots & -1 \\ N(N-1) & 0 & \cdots & 0 & -N(N-1) & 0 & \cdots & 0 \\ N - N^2 & N & \cdots & N & N - N^2 & N & \cdots & N \end{bmatrix}.$$

Therefore, the BLUE-OS in the symmetric exponential network delays case is given by

$$\begin{aligned}
\hat{\Phi}_{\text{BLUE-OS}}^{\text{S}} &= \begin{bmatrix} \hat{\tau}_{\text{BLUE-OS}}^{\text{S}} \\ \hat{\phi}_{\text{BLUE-OS}}^{\text{S}} \\ \hat{\lambda}_{\text{BLUE-OS}}^{\text{S}} \end{bmatrix} \\
&= \frac{1}{2N(N-1)} \begin{bmatrix} (N^2-1)U_{(1)} - \sum_{k=2}^N U_{(k)} + (N^2-1)V_{(1)} - \sum_{k=2}^N V_{(k)} \\ N(N-1)U_{(1)} - N(N-1)V_{(1)} \\ (N-N^2)U_{(1)} + \sum_{k=2}^N U_{(k)} + (N-N^2)V_{(1)} + \sum_{k=2}^N V_{(k)} \end{bmatrix} \\
&= \frac{1}{2(N-1)} \begin{bmatrix} N(U_{(1)} + V_{(1)}) - (\bar{U} + \bar{V}) \\ (N-1)(U_{(1)} - V_{(1)}) \\ N\{(\bar{U} + \bar{V}) - (U_{(1)} + V_{(1)})\} \end{bmatrix}, \tag{2.3}
\end{aligned}$$

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 (2.1).

2. 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 of paramount importance. Let $\Phi_{\text{BLUE-OS}}^{\text{A}} \triangleq [\tau \ \phi \ \alpha \ \beta]^T$, then the linear model based on the ordered observations

can be expressed as

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

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 $\Phi_{\text{BLUE-OS}}^{\text{A}}$ is the 4×1 vector of unknown parameters. Since the model has been shown to be linear in terms of the ordered observations, the BLUE-OS is now given by

$$\hat{\Phi}_{\text{BLUE-OS}}^{\text{A}} = (\mathbf{Q}^T \mathbf{C}_{\mathbf{z}}^{-1} \mathbf{Q})^{-1} \mathbf{Q}^T \mathbf{C}_{\mathbf{z}}^{-1} \mathbf{z},$$

where $\mathbf{C}_{\mathbf{z}}$ is the joint covariance matrix for $U_{(k)}$ and $V_{(k)}$. Due to their mutual independence, $\mathbf{C}_{\mathbf{z}}$ can be expressed as

$$\mathbf{C}_{\mathbf{z}} = \begin{bmatrix} \alpha^2 \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \beta^2 \mathbf{C} \end{bmatrix},$$

and its inverse can be written as

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

Based on the above expression, it follows that

$$\mathbf{Q}^T \mathbf{C}_z^{-1} \mathbf{Q} = \begin{bmatrix} (\alpha^{-2} + \beta^{-2})N^2 & (\alpha^{-2} - \beta^{-2})N^2 & \alpha^{-2}N & \beta^{-2}N \\ (\alpha^{-2} - \beta^{-2})N^2 & (\alpha^{-2} + \beta^{-2})N^2 & \alpha^{-2}N & -\beta^{-2}N \\ \alpha^{-2}N & \alpha^{-2}N & \alpha^{-2}N & 0 \\ \beta^{-2}N & -\beta^{-2}N & 0 & \beta^{-2}N \end{bmatrix}$$

and its inverse takes the form

$$(\mathbf{Q}^T \mathbf{C}_z^{-1} \mathbf{Q})^{-1} = \frac{1}{2N(N-1)} \begin{bmatrix} \frac{1}{2}(\alpha^2 + \beta^2) & \frac{1}{2}(\alpha^2 - \beta^2) & -\alpha^2 & -\beta^2 \\ \frac{1}{2}(\alpha^2 - \beta^2) & \frac{1}{2}(\alpha^2 + \beta^2) & -\alpha^2 & \beta^2 \\ -\alpha^2 & -\alpha^2 & 2N\alpha^2 & 0 \\ -\beta^2 & \beta^2 & 0 & 2N\beta^2 \end{bmatrix} \quad (2.4)$$

Consequently,

$$(\mathbf{Q}^T \mathbf{C}_z^{-1} \mathbf{Q})^{-1} \mathbf{Q}^T \mathbf{C}_z^{-1} = \frac{1}{2N(N-1)} \begin{bmatrix} \mathbf{A} & \mathbf{B} \end{bmatrix},$$

where the matrices \mathbf{A} and \mathbf{B} are defined as

$$\mathbf{A} = \begin{bmatrix} \frac{N^2}{\alpha^2} \left(\frac{\alpha^2 + \beta^2}{2} + \frac{\alpha^2 - \beta^2}{2} \right) - 1 & -1 & \cdots & -1 \\ \frac{N^2}{\alpha^2} \left(\frac{\alpha^2 - \beta^2}{2} + \frac{\alpha^2 + \beta^2}{2} \right) - 1 & -1 & \cdots & -1 \\ 2N - 2N^2 & 2N & \cdots & 2N \\ 0 & 0 & \cdots & 0 \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} \frac{N^2}{\beta^2} \left(\frac{\alpha^2 + \beta^2}{2} - \frac{\alpha^2 - \beta^2}{2} \right) - 1 & -1 & \cdots & -1 \\ \frac{N^2}{\beta^2} \left(\frac{\alpha^2 - \beta^2}{2} - \frac{\alpha^2 + \beta^2}{2} \right) + 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ 2N - 2N^2 & 2N & \cdots & 2N \end{bmatrix}.$$

It follows from the above equations that

$$(\mathbf{Q}^T \mathbf{C}_z^{-1} \mathbf{Q})^{-1} \mathbf{Q}^T \mathbf{C}_z^{-1} = \frac{1}{2N(N-1)} \begin{bmatrix} N^2 - 1 & -1 & \cdots & -1 & N^2 - 1 & -1 & \cdots & -1 \\ N^2 - 1 & -1 & \cdots & -1 & -(N^2 - 1) & 1 & \cdots & 1 \\ 2N - 2N^2 & 2N & \cdots & 2N & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 2N - 2N^2 & 2N & \cdots & 2N \end{bmatrix},$$

which implies

$$\begin{aligned} \hat{\Phi}_{\text{BLUE-OS}}^{\mathbf{A}} &= \begin{bmatrix} \hat{\tau}_{\text{BLUE-OS}}^{\mathbf{A}} \\ \hat{\phi}_{\text{BLUE-OS}}^{\mathbf{A}} \\ \hat{\alpha}_{\text{BLUE-OS}}^{\mathbf{A}} \\ \hat{\beta}_{\text{BLUE-OS}}^{\mathbf{A}} \end{bmatrix} \\ &= \frac{1}{2N(N-1)} \begin{bmatrix} (N^2 - 1)U_{(1)} - \sum_{k=2}^N U_{(k)} + (N^2 - 1)V_{(1)} - \sum_{k=2}^N V_{(k)} \\ (N^2 - 1)U_{(1)} - \sum_{k=2}^N U_{(k)} - (N^2 - 1)V_{(1)} + \sum_{k=2}^N V_{(k)} \\ 2NU_{(1)} - 2N^2U_{(1)} + 2N \sum_{k=2}^N U_{(k)} \\ 2NV_{(1)} - 2N^2V_{(1)} + 2N \sum_{k=2}^N V_{(k)} \end{bmatrix} \\ &= \frac{1}{2(N-1)} \begin{bmatrix} N(U_{(1)} + V_{(1)}) - (\bar{U} + \bar{V}) \\ N(U_{(1)} - V_{(1)}) - (\bar{U} - \bar{V}) \\ 2N(\bar{U} - U_{(1)}) \\ 2N(\bar{V} - V_{(1)}) \end{bmatrix}. \end{aligned} \quad (2.5)$$

C. Minimum Variance Unbiased Estimation (MVUE)

In parameter estimation, very often the ultimate goal is to find the estimator that achieves the minimum MSE, 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 minimum MSE 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.

The “turn-the-crank” procedure to derive the MVUE in estimation theory is based on the Rao-Blackwell-Lehmann-Scheffé theorem. First, the likelihood function should be factored according to Neymann-Fisher factorization theorem yielding the sufficient statistics \mathbf{T} . Then, it should be determined if the sufficient statistics are complete. Finally, either for any unbiased estimator $\check{\theta}$, $\hat{\theta} = E[\check{\theta}|\mathbf{T}]$ should be evaluated, or a function $g(\mathbf{T})$ of the sufficient statistics should be found such that $\hat{\theta} = g(\mathbf{T})$ is an unbiased estimator, producing $\hat{\theta}$ as the MVUE. The approach that we will follow next relies on similar steps.

1. 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

$$L(\tau, \phi, \alpha, \beta) = \alpha^{-N} \exp \left[-\frac{1}{\alpha} \sum_{k=1}^N \{U_k - \tau - \phi\} \right] \cdot \beta^{-N} \exp \left[-\frac{1}{\beta} \sum_{k=1}^N \{V_k - \tau + \phi\} \right] \cdot I [U_{(1)} - \tau - \phi] \cdot I [V_{(1)} - \tau + \phi], \quad (2.6)$$

where $I[\cdot]$ denotes the unit step function. Exploiting the fact that the raw sample mean and the ordered sample mean are actually the same, (2.6) can be factored as

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

where

$$g_1 \left(\sum_{k=1}^N U_{(k)}, \tau, \phi, \alpha \right) = \alpha^{-N} e^{-\frac{1}{\alpha} \sum_{k=1}^N (U_{(k)} - \tau - \phi)}, \quad g_3(U_{(1)}, \tau, \phi) = I[U_{(1)} - \tau - \phi],$$

$$g_2 \left(\sum_{k=1}^N V_{(k)}, \tau, \phi, \beta \right) = \beta^{-N} e^{-\frac{1}{\beta} \sum_{k=1}^N (V_{(k)} - \tau + \phi)}, \quad g_4(V_{(1)}, \tau, \phi) = I[V_{(1)} - \tau + \phi],$$

$$h_1(U_k, V_k) = 1.$$

In the above relations, $h_1(U_k, V_k)$ is independent of the unknown vector parameter $\Phi_{\text{MVUE}}^{\mathbf{A}-\mathbf{U}} = [\tau \ \phi \ \alpha \ \beta]^T$, whereas $g_1(\sum_{k=1}^N U_{(k)}, \tau, \phi, \alpha)$, $g_2(\sum_{k=1}^N V_{(k)}, \tau, \phi, \beta)$, $g_3(U_{(1)}, \tau, \phi)$ and $g_4(V_{(1)}, \tau, \phi)$ 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 $\Phi_{\text{MVUE}}^{\mathbf{A}-\mathbf{U}}$.

Since $\dim(\mathbf{T}) = \dim(\Phi_{\text{MVUE}}^{\mathbf{A}-\mathbf{U}})$, it is easier to determine the MVUE directly from \mathbf{T} without having to evaluate $E[\check{\Phi}_{\text{MVUE}}^{\mathbf{A}-\mathbf{U}}|\mathbf{T}]$ by finding a 4×1 vector function $\hat{\Phi}_{\text{MVUE}}^{\mathbf{A}-\mathbf{U}}$ such that $E[\hat{\Phi}_{\text{MVUE}}^{\mathbf{A}-\mathbf{U}}] = \Phi_{\text{MVUE}}^{\mathbf{A}-\mathbf{U}}$, 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 $U_{(1)}, U_{(2)}, \dots, U_{(N)}$ is given by

$$p(U_{(1)}, U_{(2)}, \dots, U_{(N)}) = N! \alpha^{-N} e^{-\frac{1}{\alpha} \sum_{k=1}^N \{U_{(k)} - \tau - \phi\}} \cdot \prod_{k=1}^N I[U_{(k)} - \tau - \phi], \quad (2.7)$$

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

$$z_k = (N - k + 1) (U_{(k)} - U_{(k-1)}), \quad k = 1, 2, \dots, N,$$

where $U_{(0)} = \tau + \phi$. Since $\sum_{k=1}^N (U_{(k)} - \tau - \phi) = \sum_{k=1}^N z_k$ and the Jacobian of the transformation is $N!$, a substitution in (2.7) reveals that

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

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, \dots, z_N is independent of z_1 (and hence of $U_{(1)}$, since $z_1 = N(U_{(1)} - \tau - \phi)$), $\sum_{k=1}^N (U_{(k)} - U_{(1)}) \sim \Gamma(N - 1, \alpha)$ 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 $\Phi_{\mathbf{M}\mathbf{V}\mathbf{U}\mathbf{E}}^{\mathbf{A}-\mathbf{U}}$ because the sufficient statistics are unique within one-to-one transformations [27]. Consequently, \mathbf{T}' comprises of four independent random variables, that in terms of the three-parameter Gamma distribution assume the distributions:

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

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 $E[g(\mathbf{T}')] = \Phi_{\text{MVUE}}^{\mathbf{A}-\mathbf{U}}$. Suppose that there exists another function h for which $E[h(\mathbf{T}')] = \Phi_{\text{MVUE}}^{\mathbf{A}-\mathbf{U}}$ is also true. Then,

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

where $\pi(\mathbf{T}') \triangleq g(\mathbf{T}') - h(\mathbf{T}')$ and the expectation is taken with respect to $p(\mathbf{T}'; \Phi_{\text{MVUE}}^{\mathbf{A}-\mathbf{U}})$.

As a result,

$$\int \int \int \int_{R_{U(1), V(1)}} \pi(r, U(1), s, V(1)) \cdot \frac{\alpha^{-(N-1)}}{\Gamma(N-1)} r^{N-2} e^{-\frac{r}{\alpha}} \cdot \frac{N}{\alpha} e^{-\frac{N}{\alpha}\{U(1) - \tau - \phi\}} \cdot \frac{\beta^{-(N-1)}}{\Gamma(N-1)} s^{N-2} e^{-\frac{s}{\beta}} \cdot \frac{N}{\beta} e^{-\frac{N}{\beta}\{V(1) - \tau + \phi\}} dr dU(1) ds dV(1) = 0 \quad \forall \Phi_{\text{MVUE}}^{\mathbf{A}-\mathbf{U}}$$

where $R_{U(1), V(1)}$ is the region defined by $I[U(1) - \tau - \phi]$ and $I[V(1) - \tau - \phi]$. The above relation can be expressed as

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \pi(r, U(1), s, V(1)) \cdot r^{N-2} s^{N-2} e^{-\left\{\frac{r}{\alpha} + \frac{NU(1)}{\alpha} + \frac{s}{\beta} + \frac{NV(1)}{\beta}\right\}} dr dU(1) ds dV(1) = 0 \quad \forall \Phi_{\text{MVUE}}^{\mathbf{A}-\mathbf{U}}$$

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 $\Phi_{\text{MVUE}}^{\mathbf{A}-\mathbf{U}}$ 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 [28] which states that *if \mathbf{T} , taking values in \mathfrak{R}^k , is sufficient for $\Phi_{\text{MVUE}}^{\mathbf{A}-\mathbf{U}}$ and boundedly complete, then \mathbf{T} is minimal sufficient.*

What remains is finding an unbiased estimator for $\Phi_{\text{MVUE}}^{\mathbf{A}-\mathbf{U}}$ as a function of \mathbf{T} , which is the MVUE according to the Rao-Blackwell-Lehmann-Scheffé theorem. At

first, it may seem difficult to find four unbiased functions of \mathbf{T} for each of τ , ϕ , α and β just by inspection. But note that BLUE-OS $\hat{\Phi}_{BLUE-OS}^A$ in (2.5) is also an unbiased function of \mathbf{T} . Hence, it is concluded that the BLUE-OS is also the MVUE.

$$\hat{\Phi}_{MVUE}^{A-U} = \begin{bmatrix} \hat{\tau}_{MVUE}^{A-U} \\ \hat{\phi}_{MVUE}^{A-U} \\ \hat{\alpha}_{MVUE}^{A-U} \\ \hat{\beta}_{MVUE}^{A-U} \end{bmatrix} = \frac{1}{2(N-1)} \begin{bmatrix} N(U_{(1)} + V_{(1)}) - (\bar{U} + \bar{V}) \\ N(U_{(1)} - V_{(1)}) - (\bar{U} - \bar{V}) \\ 2N(\bar{U} - U_{(1)}) \\ 2N(\bar{V} - V_{(1)}) \end{bmatrix}. \quad (2.8)$$

The covariance matrix of this estimator is given by (2.4) and hence minimum variances of the clock offsets, fixed and mean delay parameters are given by its diagonal elements, whereas the total mean square error for the vector parameter $\hat{\Phi}_{MVUE}^{A-U}$ is the trace of this matrix.

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

$$\hat{\phi}_{MVUE}^{A-U} = \frac{1}{N-1} \left[N \frac{U_{(1)} - V_{(1)}}{2} - \frac{\bar{U} - \bar{V}}{2} \right], \quad (2.9)$$

and its variance or MSE is written as

$$\text{var} \left(\hat{\phi}_{MVUE}^{A-U} \right) = \frac{1}{4N(N-1)} (\alpha^2 + \beta^2).$$

Similarly, the MVUE of the fixed delay τ and mean link delays α and β are the same as in (2.5). For the sake of completion, the MVUE is also given when α and β are known. It is straightforward to see from (2.6) that $U_{(1)}$ and $V_{(1)}$ are the complete minimal sufficient statistic for estimating τ and ϕ . The only unbiased functions of

$\{U_{(1)}, V_{(1)}\}$ yielding $\Phi_{\text{MVUE}}^{\text{A-K}}$ are

$$\hat{\Phi}_{\text{MVUE}}^{\text{A-K}} = \begin{bmatrix} \hat{\tau}_{\text{MVUE}}^{\text{A-K}} \\ \hat{\phi}_{\text{MVUE}}^{\text{A-K}} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} (U_{(1)} - \frac{\alpha}{N}) + (V_{(1)} - \frac{\beta}{N}) \\ (U_{(1)} - \frac{\alpha}{N}) - (V_{(1)} - \frac{\beta}{N}) \end{bmatrix}. \quad (2.10)$$

2. 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

$$L(\tau, \phi, \lambda) = \lambda^{-2N} \exp \left[-\frac{1}{\lambda} \sum_{k=1}^N \{U_k + V_k - 2\tau\} \right] \cdot I[U_{(1)} - \tau - \phi] \cdot I[V_{(1)} - \tau + \phi]. \quad (2.11)$$

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 $\Phi_{\text{MVUE}}^{\text{S-U}} = [\tau \ \phi \ \lambda]^T$. But then they have already generated *two* unbiased clock offset estimators, given by (2.3) and (2.5). Naturally, this question arises: since the same sufficient statistics have been proved complete, how can they yield two unbiased estimators? The answer to this lies in the consistency of science when we note that (2.11) can be factored as

$$L(\tau, \phi, \alpha, \beta) = g_1 \left(\sum_{k=1}^N U_{(k)}, \sum_{k=1}^N V_{(k)}, \tau, \lambda \right) g_2(U_{(1)}, \tau, \phi) g_3(V_{(1)}, \tau, \phi) h_1(U_k, V_k),$$

where

$$g_1 \left(\sum_{k=1}^N U_{(k)}, \sum_{k=1}^N V_{(k)}, \tau, \lambda \right) = \lambda^{-2N} \exp \left[-\frac{1}{\lambda} \sum_{k=1}^N \{U_k + V_k - 2\tau\} \right], \\ g_2(U_{(1)}, \tau, \phi) = I[U_{(1)} - \tau - \phi], \quad g_3(V_{(1)}, \tau, \phi) = I[V_{(1)} - \tau + \phi], \quad h_1(U_k, V_k) = 1.$$

It turns out that $\mathbf{T} = \{\sum_{k=1}^N (U_k + V_k), U_{(1)}, V_{(1)}\}$ is the actual minimal sufficient statistic instead of $\{\sum_{k=1}^N U_k, U_{(1)}, \sum_{k=1}^N V_k, V_{(1)}\}$ according to Neymann-Fisher Factorization theorem. Consequently, the clock offset estimator in (2.5) is not even a

choice to consider for not being a function of \mathbf{T} .

Now proceeding similarly as before, $\sum_{k=1}^N (U_k + V_k)$ is dependent on both $U_{(1)}$ and $V_{(1)}$. As a result, \mathbf{T} can be transformed into $\mathbf{T}' = \{\sum_{k=1}^N (U_k - U_{(1)} + V_k - V_{(1)}), U_{(1)}, V_{(1)}\}$. It is evident from the reasoning in the last subsection that $\sum_{k=1}^N (U_k - U_{(1)} + V_k - V_{(1)})$ is Gamma distributed with parameters $(2(N - 1), \lambda)$. Hence, \mathbf{T}' is a combination of three independent random variables, which in terms of the three parameter Gamma distribution assume the distributions

$$r = \sum_{k=1}^N (U_k - U_{(1)} + V_k - V_{(1)}) \sim \Gamma(2(N - 1), \lambda, 0),$$

$$U_{(1)} \sim \Gamma(1, \lambda/N, \tau + \phi), \quad V_{(1)} \sim \Gamma(1, \lambda/N, \tau - \phi).$$

Next, defining $g(\mathbf{T}')$ and $h(\mathbf{T}')$ as functions of \mathbf{T}' such that $E[g(\mathbf{T}')] = E[h(\mathbf{T}')] = \Phi_{\text{MVUE}}^{\text{S-U}}$,

$$E[g(\mathbf{T}') - h(\mathbf{T}')] = E[\pi(\mathbf{T}')] = 0 \quad \forall \Phi_{\text{MVUE}}^{\text{S-U}}$$

where the expectation is taken with respect to $p_{\mathbf{T}'}(\mathbf{T}'; \Phi_{\text{MVUE}}^{\text{S-U}})$. As a result, since the domain of r and s are controlled by $U_{(1)}$ and $V_{(1)}$, respectively

$$\int \int \int_{R_{U_{(1)}, V_{(1)}}} \pi(r, U_{(1)}, V_{(1)}) \cdot \frac{\lambda^{-\{2(N-1)\}}}{\Gamma[2(N-1)]} r^{2N-3} e^{-\frac{r}{\lambda}} \cdot \left(\frac{N}{\lambda}\right)^2 e^{-\frac{N}{\lambda}\{U_{(1)}+V_{(1)}-2\tau\}}.$$

$$dr dU_{(1)} dV_{(1)} = 0 \quad \forall \Phi_{\text{MVUE}}^{\text{S-U}}$$

where $R_{U_{(1)}, V_{(1)}}$ is the region defined by $I[U_{(1)} - \tau - \phi]$ and $I[V_{(1)} - \tau - \phi]$. It follows that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \pi(r, U_{(1)}, V_{(1)}) \cdot r^{2N-3} e^{-\frac{N}{\lambda}\{\frac{r}{N}+U_{(1)}+V_{(1)}\}} dr dU_{(1)} dV_{(1)} = 0 \quad \forall \Phi_{\text{MVUE}}^{\text{S-U}}$$

From the uniqueness theorem for the two-sided Laplace transform, it follows that $\pi(\mathbf{T}') = 0$ almost everywhere, resulting in the completeness of \mathbf{T}' , or equivalently \mathbf{T} . Hence, \mathbf{T} is also the minimal sufficient statistics from Bahadur's theorem and the

MVUE is the same as $\hat{\Phi}_{\text{BLUE-OS}}^{\text{S}}$ in (2.3) expressed as

$$\hat{\Phi}_{\text{MVUE}}^{\text{S-U}} = \begin{bmatrix} \hat{\tau}_{\text{MVUE}}^{\text{S-U}} \\ \hat{\phi}_{\text{MVUE}}^{\text{S-U}} \\ \hat{\lambda}_{\text{MVUE}}^{\text{S-U}} \end{bmatrix} = \frac{1}{2(N-1)} \begin{bmatrix} N(U_{(1)} + V_{(1)}) - (\bar{U} + \bar{V}) \\ (N-1)(U_{(1)} - V_{(1)}) \\ N\{(\bar{U} + \bar{V}) - (U_{(1)} + V_{(1)})\} \end{bmatrix}. \quad (2.12)$$

The covariance matrix of this estimator is given by (2.2) and the diagonal elements represent the variance of each individual unknown parameter, whereas the trace of this matrix is the total mean square error or variance for the vector parameter $\Phi_{\text{MVUE}}^{\text{S-U}}$.

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

$$\hat{\phi}_{\text{MVUE}}^{\text{S-U}} = \frac{U_{(1)} - V_{(1)}}{2}, \quad (2.13)$$

and its variance or MSE is

$$\text{var}\left(\hat{\phi}_{\text{MVUE}}^{\text{S-U}}\right) = \frac{\lambda^2}{2N^2}.$$

Furthermore, the MVUEs for the fixed delay τ and mean link delay λ under the symmetric assumption match the ones in (2.3). Finally, following a similar procedure, when λ is known, the sufficient statistics are $U_{(1)}$ and $V_{(1)}$ and the MVUE is

$$\hat{\Phi}_{\text{MVUE}}^{\text{S-K}} = \begin{bmatrix} \hat{\tau}_{\text{MVUE}}^{\text{S-K}} \\ \hat{\phi}_{\text{MVUE}}^{\text{S-K}} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} U_{(1)} + V_{(1)} \\ U_{(1)} - V_{(1)} \end{bmatrix}. \quad (2.14)$$

D. Explanatory Remarks

Summarizing the results derived so far, Tables I and II show 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 equations (2.8), (2.10), (2.12) and (2.14).

Table I. The MVUE for the clock offset for known mean link delay

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

Table II. The MVUE for the clock offset for unknown mean link delay

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

It is evident from Tables I and II that in practical scenarios where the means of the exponentially distributed delays are unknown, the MVUE is given by (2.9) or (2.13) 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 (2.1) with the MVUE in (2.9) as follows:

$$MSE(\phi_{MVUE}^{A-U}) = \frac{1}{4N(N-1)} (\alpha^2 + \beta^2) \quad (2.15)$$

$$\begin{aligned} MSE(\phi_{MLE}) &= \frac{1}{4N^2} (\alpha^2 + \beta^2) + \frac{1}{4N^2} (\alpha - \beta)^2 \\ &= \frac{1}{2N^2} (\alpha^2 + \beta^2 - \alpha\beta). \end{aligned} \quad (2.16)$$

Notice that though ϕ_{MLE} is biased in the most realistic setting, i.e., asymmetric unknown mean link delays, in accordance with (2.15) and (2.16), it outperforms the MVUE under the condition

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

which can be expressed equivalently as:

$$\frac{N}{2} - 1 < \frac{\alpha\beta}{(\alpha - \beta)^2} \triangleq f(\alpha, \beta). \quad (2.17)$$

The above relations bring into attention a number of remarks. First, (2.17) 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 (2.17). 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 does 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 (2.15) and (2.16) 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, (2.17) 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 (2.8) and (2.12) and plugging in (2.17) might be a good

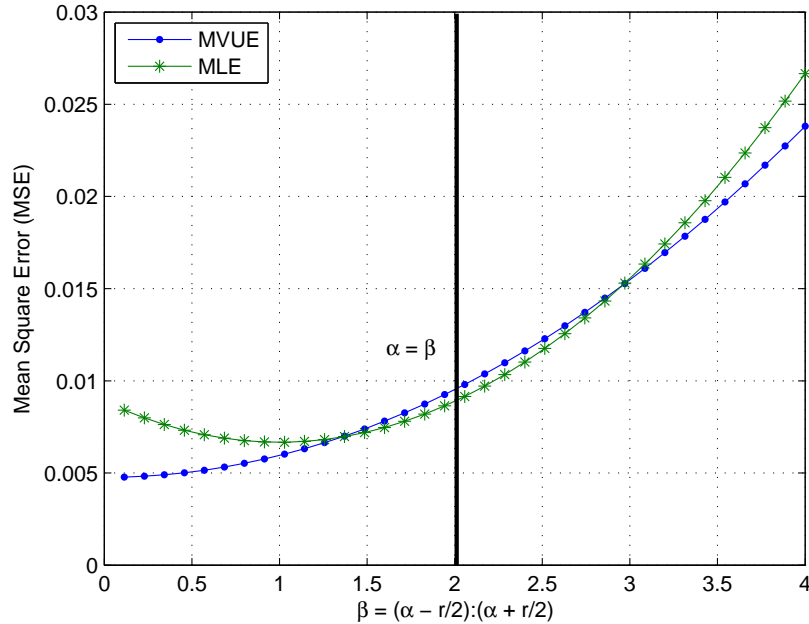


Fig. 2. Mean square error of the MLE (2.1) and the MVUE (2.9) for asymmetric unknown delays with constant $N = 15$.

idea for adaptively selecting between the MVUE and MLE as

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

However, since $f(\hat{\alpha}, \hat{\beta})$ processes nonlinearly the estimates, a considerable amplification of estimation errors occurs which affects the quality of the resultant $f(\hat{\alpha}, \hat{\beta})$. In other words, even having access to $\hat{\alpha}_{MLE}$ and $\hat{\beta}_{MLE}$ does 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. These findings are very important in the context of WSNs where energy resources are limited and the number of synchronization packet exchanges is rather small. Even in the traditional centralized or ad-hoc networks,

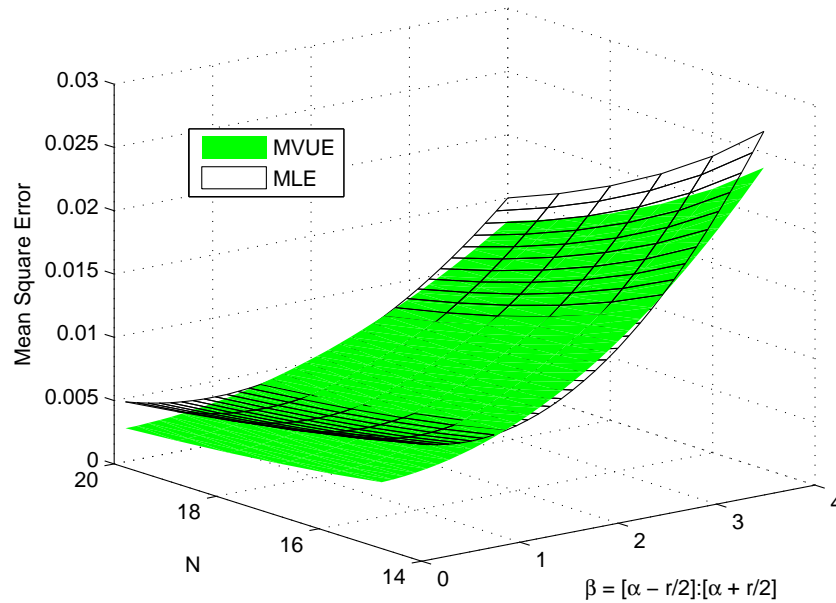


Fig. 3. Mean square error of the MLE (2.1) and the MVUE (2.9) for asymmetric unknown delays with different values of N .

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 an excellent textbook 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 justifies the reason why it has been the most widely used estimator to date 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 (2.8) and (2.12), 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 and usually there is no provision inside the network to provide end-systems with information about the current status of the network. For example, packet delay statistics are important in examining the performance and reliability of the Internet, but it has no mechanism for providing 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 variable delays are also important in other areas such as continuous-media applications, e.g., audio and video applications need to absorb the delay jitter perceived at the receiver for smooth playout of the original stream (see [29] and [30]). For better performance of such applications, determining the correct amount of buffering, and the reconstruction of the original timing plays a significant role.

E. Summary

The MLE of the clock offset from timing message exchanges between two clocks were derived in [18] when the fixed delays are symmetric and the variable delays in each direction have an exponential distribution with an unknown mean. In this chapter, the BLUE-OS of the clock offset between two nodes are derived assuming both symmetric and asymmetric exponential network delays. The Rao-Blackwell-Lehmann-Scheffé theorem is then exploited to obtain the MVUE for the clock offset

and 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 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.

CHAPTER III

CLOCK OFFSET AND SKEW ESTIMATION*

Advancing one step further, we now turn our attention to a more accurate model defining the relationship between two clocks by the addition of clock skew. In practice, the time synchronization problem in WSNs generally involves two steps: synchronizing the nodes in the network to one common absolute time by adjusting clock phase offset (clock offset) among the nodes, and correcting the clock frequency offset (clock skew) relative to a certain standard frequency. The second step is required because the imperfections in the quartz crystal and environmental conditions induce different clocks run at slightly different frequencies. Actually, the effect of clock skew is the main reason why clock offset keeps drifting away. Hence, adjusting clock skew guarantees long-term reliability of synchronization, and therefore reduces network-wide energy consumption in synchronization procedures. Indeed, developing long-term and network-wide time synchronization protocols that are energy-efficient represents one of the key strategies for the successful deployment of long-lived WSNs. The main contributions of this chapter are as follows.

1. The MLE and the corresponding CRLB for the conventional clock offset model in a general sender-receiver protocol assuming Gaussian model for the noise are derived.

*Parts of this chapter are reprinted with permissions from “Novel Clock Phase Offset and Skew Estimation Using Two-Way Timing Message Exchanges for Wireless Sensor Networks” by Kyoung-Lae Noh, Qasim M. Chaudhari, Erchin Serpedin and Bruce Suter, *IEEE Transactions on Communications*, Volume 55, Issue 4, April 2007 Page(s): 766 - 777 and “On Maximum Likelihood Estimation of Clock Offset and Skew in Networks With Exponential Delays”, by Qasim M. Chaudhari, Erchin Serpedin and Khalid Qaraqe, *IEEE Transactions on Signal Processing*, Volume 56, Issue 4, April 2008 Page(s): 1685 - 1697.

2. The joint MLE and corresponding CRLB using a more realistic linear clock offset and skew model assuming Gaussian random delays are obtained.
3. The CRLB for the clock offset for exponential delay model is derived as a performance threshold.
4. The joint MLE for the clock offset and skew under the exponential delay model is obtained and the corresponding algorithms to find these estimators are also described in detail.

A. Gaussian Delay Model

As explained earlier, several pdf models have been proposed for random network delays. Even for an unknown delay distribution, the final error is a sum of many independent random components described in Chapter I. Exploiting the Central Limit theorem, which asserts that the pdf of the sum of a number of independent and identically distributed (iid) random variables approaches that of a Gaussian random variable, the Gaussian model in our study will be appropriate if the delays are thought to be the addition of a few such independent random processes. For example, suppose that the actual errors are uniformly distributed around a mean delay value, then the sum of just two such errors closely resembles Gaussian pdf. In addition, the Gaussian distribution for the phase offset errors is reported by a few authors, such as [10], based on the laboratory tests. In this section, we derive the MLE and the CRLB for both the clock offset only and the clock offset and skew models under Gaussian delay assumption.

1. Maximum Likelihood Clock Offset Estimation

Assuming no clock skew at this stage and utilizing the same mechanism as in Chapter II, we compute the MLE and CRLB for the clock offset using the two-way timing message exchange model. Since the set of delay observations $\{X_k\}_{k=1}^N$ and $\{Y_k\}_{k=1}^N$ are independently and normally distributed with the same mean μ and variance σ^2 based on Gaussian assumption, the likelihood function based on the observations $\{X_k\}_{k=1}^N$ and $\{Y_k\}_{k=1}^N$ is given by

$$L(\phi, \mu, \sigma^2) = (2\pi\sigma^2)^{-N} e^{-\frac{1}{2\sigma^2} \left[\sum_{k=1}^N (U_k - \tau - \phi - \mu)^2 + \sum_{k=1}^N (V_k - \tau + \phi - \mu)^2 \right]}.$$

Differentiating the log-likelihood function gives

$$\begin{aligned} \frac{\partial \ln L(\phi)}{\partial \phi} &= -\frac{1}{2\sigma^2} \left[\sum_{k=1}^N (2\phi - 2(U_k - \tau - \mu)) + \sum_{k=1}^N (2\phi + 2(V_k - \tau - \mu)) \right] \\ &= -\frac{1}{\sigma^2} \left[\sum_{k=1}^N (2\phi - (U_k - V_k)) \right]. \end{aligned} \quad (3.1)$$

Hence, the MLE of clock offset is given by

$$\hat{\phi} = \arg \max_{\phi} [\ln L(\phi)] = \frac{\sum_{k=1}^N (U_k - V_k)}{2N} = \frac{\bar{U} - \bar{V}}{2}. \quad (3.2)$$

Consequently, the MLE of clock offset can be obtained by finding the means of observations $\{U_k\}_{k=1}^N$ and $\{V_k\}_{k=1}^N$.

2. Cramer-Rao Lower Bound for Clock Offset

The regularity condition [27] holds for the given estimate since the expected value of (3.1) is 0. Thus, the CRLB for the MLE can be obtained by differentiating (3.1) w.r.t. ϕ , which gives

$$\frac{\partial^2 \ln L(\phi)}{\partial \phi^2} = -\frac{2N}{\sigma^2}.$$

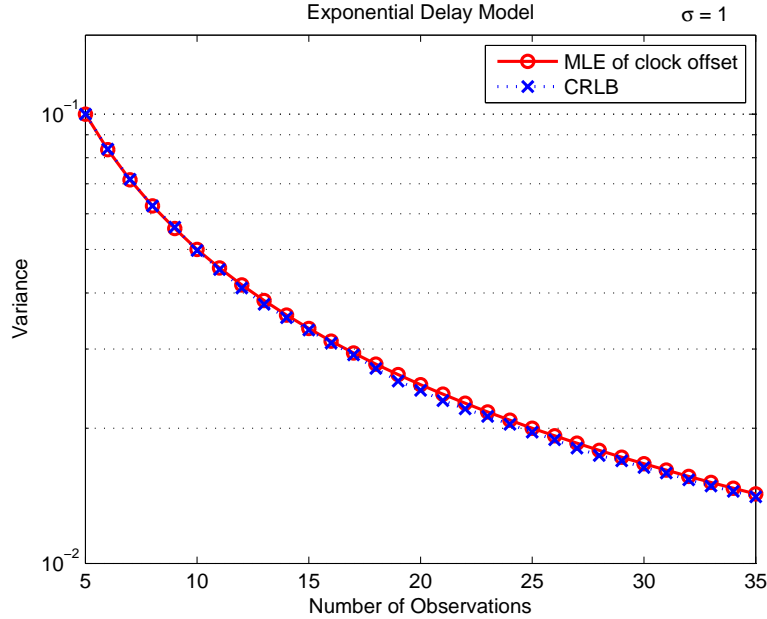


Fig. 4. CRLB and variance of the MLE of clock offset for the Gaussian delay model ($\sigma = 1$).

Hence the CRLB for the MLE is given by

$$\text{var}(\hat{\phi}) \geq -E \left[\frac{\partial^2 \ln L(\phi)}{\partial \phi^2} \right]^{-1} = \frac{\sigma^2}{2N}. \quad (3.3)$$

Fig. 4 shows the result of the computer simulation when σ is 1. It can be seen that the variance of estimate is proportional to σ^2 and inversely proportional to N .

3. Joint Maximum Likelihood Estimation of Clock Offset and Skew

Since every oscillator has its unique clock frequency, the clock offset between two nodes generally keeps increasing. Therefore, a fixed value model for clock time difference as above is not sufficient for some practical situations. Hence, estimating the difference of clock frequencies between two nodes (i.e., clock skew) increases synchronization accuracy and guarantees long-term reliability. In this section, we derive the joint MLE for clock offset and skew based on the two-way timing message exchange

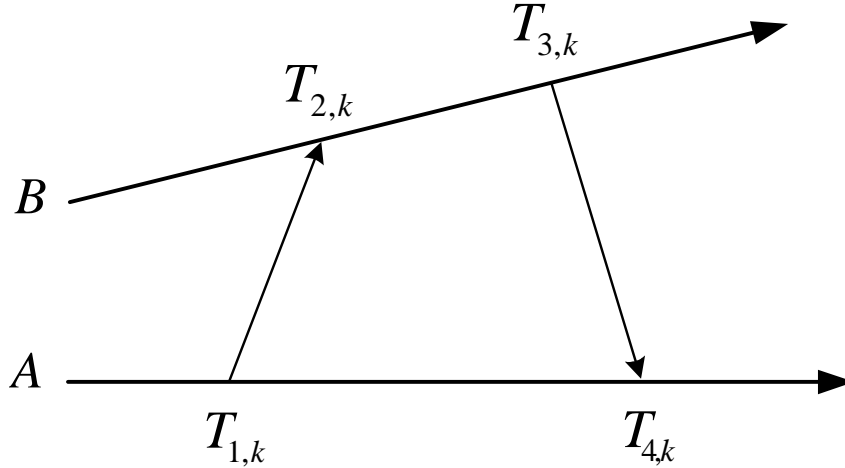


Fig. 5. Two-way timing message exchange model having clock offset and skew.

model with Gaussian delays.

The theory applied thus far for finding the MLE and CRLB for the clock offset (assuming no clock skew) can be extended to find the joint MLE and CRLB for a more general clock model. Fig. 5 shows the effect of clock offset (ϕ) and skew (ω) on timing message exchanges between two nodes. Here, timestamps in the k th message exchange $T_{1,k}$ and $T_{4,k}$ are measured by local clock of Node A, and $T_{2,k}$ and $T_{3,k}$ are measured by local clock of Node B, respectively. Node A transmits a synchronization packet, containing the level and ID of Node A and the value of timestamp $T_{1,k}$, to Node B. Node B receives it at $T_{2,k}$ and transmits an acknowledgement packet to Node A at $T_{3,k}$. This packet contains the level and ID of Node B and the value of timestamps $T_{1,k}$, $T_{2,k}$, and $T_{3,k}$. Then Node A finally receives the packet at $T_{4,k}$.

Taking $T_{1,1}$ be the reference time, the timestamp at Node B in the k th uplink message $T_{2,k}$, is given by

$$T_{2,k} = (T_{1,k} + \tau + X_k) \omega + \phi, \quad (3.4)$$

and the timestamp at Node B in the k th downlink message $T_{3,k}$, is represented by

$$T_{3,k} = (T_{4,k} - \tau - Y_k)\omega + \phi. \quad (3.5)$$

Assuming $\{X_k\}_{k=1}^N$ and $\{Y_k\}_{k=1}^N$ are zero mean independent Gaussian distributed RVs with variance σ^2 , then the joint PDF of $\mathbf{X} \triangleq \{X_k\}_{k=1}^N$ and $\mathbf{Y} \triangleq \{Y_k\}_{k=1}^N$ is given by

$$f_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y}) = (2\pi\sigma^2)^{-N} e^{-\frac{1}{2\sigma^2} \sum_{k=1}^N \left[\left(\frac{T_{2,k}-\phi}{\omega} - T_{1,k}-\tau \right)^2 + \left(T_{4,k}-\tau - \frac{T_{3,k}-\phi}{\omega} \right)^2 \right]}.$$

Further assuming that the fixed portion of delay τ is known and $\omega' \triangleq 1/\omega$, then the likelihood function for $(\phi, \omega', \sigma^2)$, based on observations $\{T_{1,k}\}_{k=1}^N$, $\{T_{2,k}\}_{k=1}^N$, $\{T_{3,k}\}_{k=1}^N$, and $\{T_{4,k}\}_{k=1}^N$, is given by

$$L(\phi, \omega', \sigma^2) = (2\pi\sigma^2)^{-N} e^{-\frac{1}{2\sigma^2} \sum_{k=1}^N \left\{ [\omega'(T_{2,k}-\phi)-(T_{1,k}+\tau)]^2 + [\omega'(\phi-T_{3,k})+(T_{4,k}-\tau)]^2 \right\}}.$$

Differentiating the log-likelihood function with respect to ϕ gives

$$\frac{\partial \ln L(\phi, \omega', \sigma^2)}{\partial \phi} = -\frac{1}{\sigma^2} \sum_{k=1}^N \left[\omega'^2 (2\phi - T_{2,k} - T_{3,k}) + \omega' (T_{1,k} + T_{4,k}) \right]. \quad (3.6)$$

Hence, in the given clock skew model, the joint MLE of clock offset $\hat{\phi}$ can be obtained by

$$\hat{\phi} = \frac{\sum_{k=1}^N [\hat{\omega}' (T_{2,k} + T_{3,k}) - (T_{1,k} + T_{4,k})]}{2N\hat{\omega}'} \quad (3.7)$$

Note that, for the clock skew model with Gaussian random delays, there is an additional term which is depending on $\hat{\omega}$, and this result reduces to (3.2) when $\hat{\omega}$ is one.

Similarly, differentiating the log-likelihood function with respect to ω' gives

$$\frac{\partial \ln L(\phi, \omega', \sigma^2)}{\partial \omega'} = -\frac{1}{\sigma^2} \left\{ \sum_{k=1}^N \omega' [(T_{2,k} - \phi)^2 + (T_{3,k} - \phi)^2] - \sum_{k=1}^N [(T_{1,k} + \tau)(T_{2,k} - \phi) + (T_{4,k} - \tau)(T_{3,k} - \phi)] \right\}. \quad (3.8)$$

Thus, the estimate $\hat{\omega}'$ maximizing the log-likelihood function is given by

$$\hat{\omega}' = \frac{\sum_{k=1}^N [(T_{1,k} + \tau)(T_{2,k} - \hat{\phi}) + (T_{4,k} - \tau)(T_{3,k} - \hat{\phi})]}{\sum_{k=1}^N [(T_{2,k} - \hat{\phi})^2 + (T_{3,k} - \hat{\phi})^2]}.$$

Hence, the joint MLE of clock skew $\hat{\omega}$ is given by

$$\hat{\omega} = \frac{\sum_{k=1}^N [(T_{2,k} - \hat{\phi})^2 + (T_{3,k} - \hat{\phi})^2]}{\sum_{k=1}^N [(T_{1,k} + \tau)(T_{2,k} - \hat{\phi}) + (T_{4,k} - \tau)(T_{3,k} - \hat{\phi})]}. \quad (3.9)$$

In the sequel, the joint MLE of ϕ and ω can be obtained by plugging the expression of $\hat{\phi}$ (3.7) into that of $\hat{\omega}$ (3.9), which implies

$$\hat{\omega} = \frac{\sum_{k=1}^N (T_{2,k} + T_{3,k}) - 2N\hat{\phi}}{\sum_{k=1}^N (T_{1,k} + T_{4,k})} = \frac{\sum_{k=1}^N [(T_{2,k} - \hat{\phi})^2 + (T_{3,k} - \hat{\phi})^2]}{\sum_{k=1}^N [(T_{1,k} + \tau)(T_{2,k} - \hat{\phi}) + (T_{4,k} - \tau)(T_{3,k} - \hat{\phi})]}. \quad (3.10)$$

Then (3.10) can be rewritten as

$$\begin{aligned} & \sum_{k=1}^N (T_{1,k}T_{2,k} + T_{3,k}T_{4,k} + (T_{2,k} - T_{3,k})\tau) \sum_{k=1}^N (T_{2,k} + T_{3,k}) + 2N \sum_{k=1}^N (T_{1,k} + T_{4,k})\hat{\phi}^2 \\ & - \left[\sum_{k=1}^N (T_{2,k} + T_{3,k}) \sum_{k=1}^N (T_{1,k} + T_{4,k}) + 2N \sum_{k=1}^N (T_{1,k}T_{2,k} + T_{3,k}T_{4,k} + (T_{2,k} - T_{3,k})\tau) \right] \hat{\phi} \\ & = \sum_{k=1}^N (T_{1,k} + T_{4,k}) \sum_{k=1}^N (T_{2,k}^2 + T_{3,k}^2) + 2N \sum_{k=1}^N (T_{1,k} + T_{4,k})\hat{\phi}^2 \\ & - 2 \sum_{k=1}^N (T_{2,k} + T_{3,k}) \sum_{k=1}^N (T_{1,k} + T_{4,k})\hat{\phi}. \end{aligned}$$

After some manipulations, the joint MLE of clock offset and skew is given by

$$\hat{\phi}_{GML} = \frac{\sum_{k=1}^N (T_{1,k} + T_{4,k}) \sum_{k=1}^N (T_{2,k}^2 + T_{3,k}^2) - \sum_{k=1}^N (T_{2,k} + T_{3,k})Q}{\sum_{k=1}^N (T_{2,k} + T_{3,k}) \sum_{k=1}^N (T_{1,k} + T_{4,k}) - 2NQ}, \quad (3.11)$$

$$\hat{\omega}_{GML} = \frac{-2N \left[\sum_{k=1}^N (T_{1,k} + T_{4,k}) \sum_{k=1}^N (T_{2,k}^2 + T_{3,k}^2) - Q \sum_{k=1}^N (T_{2,k} + T_{3,k}) \right]}{\sum_{k=1}^N (T_{1,k} + T_{4,k}) \left[\sum_{k=1}^N (T_{2,k} + T_{3,k}) \sum_{k=1}^N (T_{1,k} + T_{4,k}) - 2NQ \right]} + \frac{\sum_{k=1}^N (T_{2,k} + T_{3,k})}{\sum_{k=1}^N (T_{1,k} + T_{4,k})}, \quad (3.12)$$

where $Q \triangleq \sum_{k=1}^N (T_{1,k}T_{2,k} + T_{3,k}T_{4,k} + (T_{2,k} - T_{3,k})\tau)$. Note that the joint MLE depends on the value of the fixed portion of delays τ , which is assumed to be known in this section. Although it is achievable, we do not consider τ as another unknown (nuisance) parameter due to its highly nonlinear and complex results.

Finally, notice that numerous authors use a linear regression model for the relationship between the clocks of two nodes, i.e., (eliminating known τ)

$$T_{2,k} = T_{1,k}\omega + \phi + X_k, \quad k = 1, \dots, N,$$

which is a restrictive assumption, since ω has its effect on the clock of the second node during the occurrence of the message delay until the message is timestamped on its reception. This is particularly important in the networks where the message delays are large, e.g., underwater acoustic networks. We assert that using more realistic model as in (3.4) and (3.5) results in superior performance of the synchronization protocol. To show the effect on MSE by using the more realistic model, we have simulated the MSE of two clock offset estimators, one of which is derived based on the correct model while the other is the standard least squares solution based on

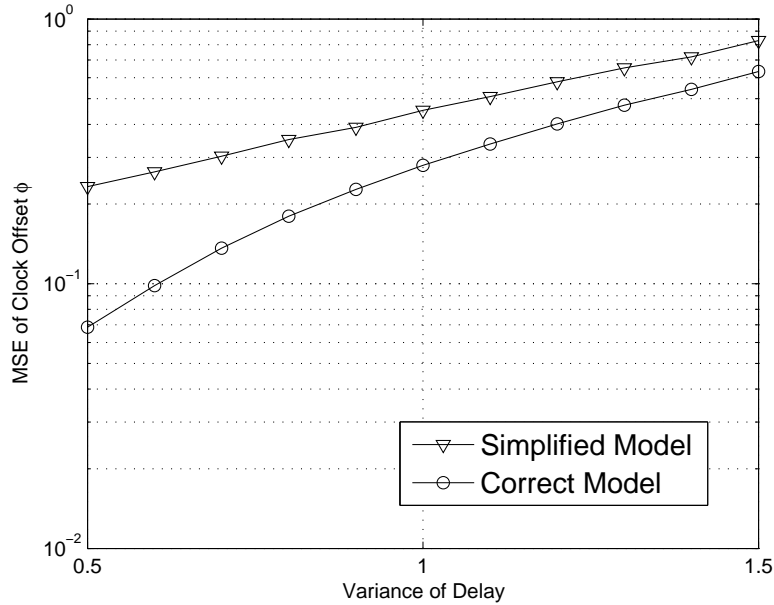


Fig. 6. MSE of clock offset estimate $\hat{\phi}$ as a function of variance of delay.

linear regression and are shown in Fig. 6. It is clear that using the more realistic model results in a lower MSE of the clock offset estimate.

4. Cramer-Rao Lower Bound for Clock Offset and Skew

The CRLB for the vector parameter $\boldsymbol{\theta} = [\phi, \omega]^T$ can be derived from the 2×2 Fisher information matrix $\mathbf{I}(\boldsymbol{\theta})$ by taking its inverse. From (3.6) and (3.8), the 2nd order derivatives of the log-likelihood function with respect to ϕ and ω' are found as

$$\begin{aligned} \frac{\partial^2 \ln L(\phi, \omega', \sigma^2)}{\partial \phi^2} &= -\frac{2N\omega'^2}{\sigma^2}, \\ \frac{\partial^2 \ln L(\phi, \omega', \sigma^2)}{\partial \omega'^2} &= -\frac{1}{\sigma^2} \sum_{k=1}^N [(T_{2,k} - \phi)^2 + (T_{3,k} - \phi)^2], \\ \frac{\partial^2 \ln L(\phi, \omega', \sigma^2)}{\partial \phi \omega'} &= -\frac{1}{\sigma^2} \sum_{k=1}^N (2\omega' \phi - \omega' T_{2,k} + T_{1,k} - \omega' T_{3,k} - T_{4,k}). \end{aligned}$$

Taking the negative expectations yields

$$\begin{aligned}
-E \left[\frac{\partial^2 \ln L(\phi, \omega', \sigma^2)}{\partial \phi^2} \right] &= \frac{2N\omega'^2}{\sigma^2}, \\
-E \left[\frac{\partial^2 \ln L(\phi, \omega', \sigma^2)}{\partial \omega'^2} \right] &= \frac{1}{\sigma^2} \sum_{k=1}^N E_{X_k, Y_k} \left[\frac{(X_k + T_{1,k} + \tau)^2 + (Y_k - T_{4,k} + \tau)^2}{\omega'^2} \right] \\
&\stackrel{(a)}{=} \frac{\sum_{k=1}^N ((T_{1,k} + \tau)^2 + (T_{4,k} - \tau)^2 + 2\sigma^2)}{\sigma^2 \omega'^2}, \\
-E \left[\frac{\partial^2 \ln L(\phi, \omega', \sigma^2)}{\partial \phi \omega'} \right] &= -\frac{1}{\sigma^2} \sum_{k=1}^N E_{X_k, Y_k} [2\omega' (2\phi - T_{2,k} - T_{3,k}) + T_{1,k} + T_{4,k}] \\
&\stackrel{(b)}{=} \frac{N}{\sigma^2} (\overline{T_1} + \overline{T_4}),
\end{aligned}$$

where (a) and (b) are due to $X_k = \omega'(T_{2,k} - \phi) - (T_{1,k} + \tau)$ and $Y_k = \omega'(\phi - T_{3,k}) + (T_{4,k} - \tau)$. Therefore, the Fisher information matrix becomes

$$\begin{aligned}
\mathbf{I}(\boldsymbol{\theta}) &= \begin{bmatrix} -E \left[\frac{\partial^2 \ln L(\phi, \omega', \sigma^2)}{\partial \phi^2} \right] & -E \left[\frac{\partial^2 \ln L(\phi, \omega', \sigma^2)}{\partial \phi \omega'} \right] \\ -E \left[\frac{\partial^2 \ln L(\phi, \omega', \sigma^2)}{\partial \omega' \phi} \right] & -E \left[\frac{\partial^2 \ln L(\phi, \omega', \sigma^2)}{\partial \omega'^2} \right] \end{bmatrix}, \\
&= \frac{1}{\sigma^2} \begin{bmatrix} 2N\omega'^2 & N(\overline{T_1} + \overline{T_4}) \\ N(\overline{T_1} + \overline{T_4}) & \frac{1}{\omega'^2} \sum_{k=1}^N [(T_{1,k} + \tau)^2 + (T_{4,k} - \tau)^2 + 2\sigma^2] \end{bmatrix} \quad (3.13)
\end{aligned}$$

Now the CRLB can be obtained by taking the inverse of the $[k, k]$ th element of the Fisher information matrix (i.e., $\text{var}(\hat{\theta}_k) \geq [\mathbf{I}^{-1}(\boldsymbol{\theta})]_{ii}$), and the inverse $\mathbf{I}^{-1}(\boldsymbol{\theta})$ is given by

$$\mathbf{I}^{-1}(\boldsymbol{\theta}) = \sigma^2 \begin{bmatrix} \frac{V}{\omega'^2 N [2V - N(\overline{T_1} + \overline{T_4})^2]} & \frac{-(\overline{T_1} + \overline{T_4})}{2V - N(\overline{T_1} + \overline{T_4})^2} \\ \frac{-(\overline{T_1} + \overline{T_4})}{2V - N(\overline{T_1} + \overline{T_4})^2} & \frac{2\omega'^2}{2V - N(\overline{T_1} + \overline{T_4})^2} \end{bmatrix}, \quad (3.14)$$

where $V = \sum_{k=1}^N [(T_{1,k} + \tau)^2 + (T_{4,k} - \tau)^2 + 2\sigma^2]$. Consequently, from the result in [27], the CRLBs of clock offset and skew for the Gaussian delay model are respectively

given by

$$\text{var}(\hat{\phi}^{GML}) \geq \frac{\sigma^2 \omega^2 V}{N \left[2V - N (\overline{T}_1 + \overline{T}_4)^2 \right]}, \quad (3.15)$$

$$\begin{aligned} \text{var}(\hat{\omega}^{GML}) &\geq \left(\frac{\partial \omega}{\partial \omega'} \right)^2 \cdot \frac{2\sigma^2 \omega'^2}{2V - N (\overline{T}_1 + \overline{T}_4)^2} \\ &= \frac{2\sigma^2 \omega^2}{2V - N (\overline{T}_1 + \overline{T}_4)^2}. \end{aligned} \quad (3.16)$$

B. Exponential Delay Model

A detailed justification of modeling the network delays as coming from an exponential distribution was presented in Chapter II. Since the MLE for the clock offset under exponential delays has been derived in [18], we derive the corresponding CRLB for the clock offset in the next section. Afterwards, the joint MLE for both the clock offset and skew is obtained and the corresponding algorithms for finding those estimates are also presented.

1. Cramer-Rao Lower Bound for Clock Offset

It was proven in [18] that the MLE of ϕ exists when τ is unknown and exhibits the same form as the estimator proposed in [19], which is given by

$$\hat{\phi} = \frac{U_{(1)} - V_{(1)}}{2}, \quad (3.17)$$

where N stands for the number of observations of delay measurements and the subscript (1) denotes the first order statistic of the corresponding data set. In this section, we proceed towards obtaining the CRLB for this clock offset under exponential delay model.

Note that (3.17) can be rewritten as

$$\hat{\phi} = \frac{U_{(1)} - V_{(1)}}{2} = \phi + \frac{X_{(1)} - Y_{(1)}}{2},$$

where $X_{(1)}$ and $Y_{(1)}$ denote the corresponding order statistics of $\{X_k\}_{k=1}^N$ and $\{Y_k\}_{k=1}^N$, respectively. Let $Z \triangleq X_{(1)} - Y_{(1)}$, then the pdf of Z can be found as follows. Since the order statistics $X_{(1)}$ and $Y_{(1)}$ are independent, $f_Z(z)$ can be found by transforming a joint distribution using the dummy variable $S = Y_{(1)}$. From the assumptions, the PDF of the uplink and downlink delays, X_k and Y_k , are given by

$$\begin{aligned} f_{X_k}(x) &= \frac{1}{\lambda_1} e^{-\frac{x}{\lambda_1}} & x \geq 0, \\ f_{Y_k}(y) &= \frac{1}{\lambda_2} e^{-\frac{y}{\lambda_2}} & y \geq 0. \end{aligned}$$

It is well known that the pdfs of the order statistics $X_{(1)}$ and $Y_{(1)}$ are given by

$$\begin{aligned} f_{X_{(1)}}(x) &= N(1 - F_{X_i}(x))^{N-1} f_{X_i}(x) = \frac{N}{\lambda_1} e^{-\frac{N}{\lambda_1}x} & x \geq 0, \\ f_{Y_{(1)}}(y) &= N(1 - F_{Y_i}(y))^{N-1} f_{Y_i}(y) = \frac{N}{\lambda_2} e^{-\frac{N}{\lambda_2}y} & y \geq 0. \end{aligned}$$

Since the *Jacobian* of this transformation is 1, a joint distribution of RVs Z and S is given by

$$\begin{aligned} f_{Z,S}(z, s) &= f_{X_{(1)}, Y_{(1)}}(z + s, s) = f_{X_{(1)}}(z + s) \cdot f_{Y_{(1)}}(s) \\ &= \frac{N^2}{\lambda_1 \lambda_2} e^{-\frac{N}{\lambda_1}z} e^{-N\left(\frac{\lambda_1 + \lambda_2}{\lambda_1 \lambda_2}\right)s} & z \geq -s, s \geq 0. \end{aligned} \quad (3.18)$$

Integrating (3.18) with respect to s yields

$$f_Z(z) = \begin{cases} \frac{N}{(\lambda_1 + \lambda_2)} e^{-\frac{N}{\lambda_1}z} & z > 0 \\ \frac{N}{(\lambda_1 + \lambda_2)} e^{\frac{N}{\lambda_2}z} & z < 0 \end{cases}. \quad (3.19)$$

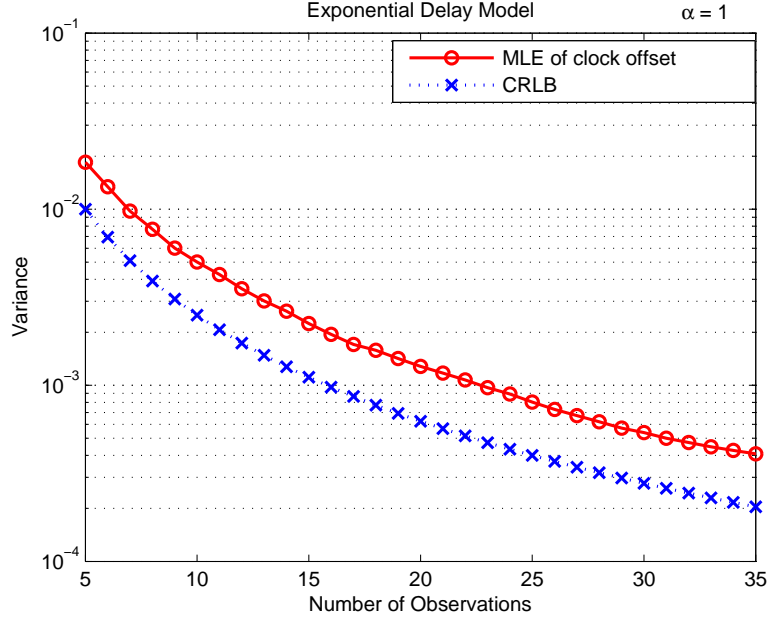


Fig. 7. CRLB and variance of the MLE of clock offset for the exponential delay model ($\alpha = 1$).

Let $W \triangleq U_{(1)} - V_{(1)}$, then the PDF of W as a function of ϕ is given by

$$f_w(w; \phi) = \begin{cases} \frac{N}{(\lambda_1 + \lambda_2)} e^{-\frac{N}{\lambda_1}(w-2\phi)} & w > 2\phi \\ \frac{N}{(\lambda_1 + \lambda_2)} e^{\frac{N}{\lambda_2}(w-2\phi)} & w < 2\phi \end{cases}. \quad (3.20)$$

Note that the estimate $\hat{\phi}$ will be biased when uplink and downlink delays are asymmetrically distributed, i.e., $\lambda_1 \neq \lambda_2$. Thus, to derive the CRLB for the estimator, the delays are assumed to be symmetric, which yields $\lambda_1 = \lambda_2 = \alpha$. Now (3.20) can be rewritten as

$$f_W(w; \phi) = \frac{N}{2\alpha} e^{-\frac{N}{\alpha}|w-2\phi|}.$$

Differentiating the logarithm of (3.20) with respect to ϕ gives

$$\frac{\partial \ln f_W(w; \phi)}{\partial \phi} = \begin{cases} \frac{2N}{\alpha} & w > 2\phi \\ -\frac{2N}{\alpha} & w < 2\phi \end{cases}, \quad (3.21)$$

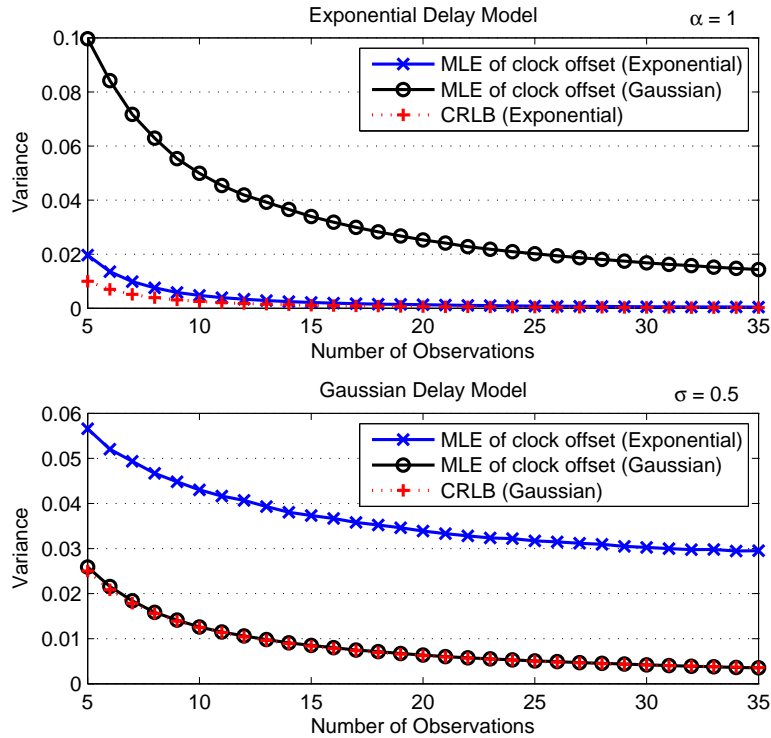


Fig. 8. Variances of both MLEs of clock offset for exponential and Gaussian delays ($\alpha = 1$ and $\sigma = 0.5$).

where the regularity condition of the CRLB holds since (3.21) is finite and the expected value of (3.21) is 0. Calculating the expected value of the square of (3.21) gives

$$E \left[\left(\frac{\partial \ln f_W(w; \phi)}{\partial \phi} \right)^2 \right] = \frac{4N^2}{\alpha^2}.$$

Therefore, the CRLB of clock offset, $\hat{\phi}$, is given by

$$\text{var}(\hat{\phi}) \geq E \left[\left(\frac{\partial \ln f_W(w; \phi)}{\partial \phi} \right)^2 \right]^{-1} = \frac{\alpha^2}{4N^2}. \quad (3.22)$$

Fig. 7 shows the simulation results corresponding to the variance and CRLB of the MLE when α is 1. It can be seen that the variance of estimate goes to zero as N increases (asymptotically efficient), and is proportional to α^2 .

In Fig. 8, the variances of both MLEs are compared in exponential and normal random delay channels, respectively. It can be seen that the performance of the ML clock offset estimator is strongly dependent on the type of random delay models.

2. Joint Maximum Likelihood Estimation of Clock Offset and Skew

From the timing message exchange model described in (3.4) and (3.5) but with exponential delays, the general form of the likelihood function is given by

$$L(\alpha, \tau, \omega, \phi) = \alpha^{-2N} \cdot \exp \left[-\frac{1}{\alpha} \left\{ \sum_{k=1}^N \frac{T_{2,k} - T_{3,k}}{\omega} - \sum_{k=1}^N (T_{1,k} - T_{4,k}) - 2Nd \right\} \right] \times \prod_{k=1}^N I \left[\frac{T_{2,k} - \phi}{\omega} - T_{1,k} - \tau \geq 0; T_{4,k} - \frac{T_{3,k} - \phi}{\omega} - \tau \geq 0 \right], \quad (3.23)$$

where the indicator function $I[\cdot]$ is defined as

$$I[\tau \geq 0] = \begin{cases} 1, & \tau \geq 0 \\ 0, & \tau < 0 \end{cases}.$$

Note that the τ is always positive since it represents the delay (fixed), while ω is also always positive because it has been realistically assumed that none of the clocks is either standing still ($\omega = 0$) or running backward ($\omega < 0$). An ideal value of $\omega = 1$ means that the clock is running at the standard rate. Also, notice that when $\omega = 1$, the MLE of clock offset ϕ was derived by [18] and takes the form

$$\hat{\phi} = \frac{1}{2} \left[\min_{1 \leq k \leq N} (T_{2,k} - T_{1,k}) - \min_{1 \leq k \leq N} (T_{4,k} - T_{3,k}) \right]. \quad (3.24)$$

From here onwards, without losing any generalization, we will assume that α is known. This is because even if α is unknown, due to the form of the reduced likelihood function $L(\tau, \omega, \phi)$ as shown in [18], the MLE $(\hat{\tau}, \hat{\omega}, \hat{\phi})$ remains the same. When $\omega \neq 1$, in maximizing the likelihood for this model over the set $\Theta = \{(\tau, \phi, \omega) :$

$\tau > 0, -\infty < \phi < \infty, \omega > 0\}$, four different cases will be considered:

Case I: τ known, ϕ known.

Case II: τ unknown, ϕ known.

Case III: τ known, ϕ unknown.

Case IV: τ unknown, ϕ unknown.

An important remark needs to be mentioned here. A preliminary examination of Cases I and II (i.e., when ϕ is known) is necessary because it gives insight into the shape of the support region over which the likelihood function is nonzero. As it is the case with exponential models, the MLEs for the location parameters will be found by taking effectively into account the boundary conditions. For the first two cases, the support of the likelihood region is a 2-D region and it is relatively easier to find the parameters on the boundary maximizing the likelihood function. Finding the MLEs for Cases III and IV (i.e., when ϕ is unknown) requires the visualization of the likelihood function support region in 3-D and getting a somewhat primitive knowledge of the 2-D support region for the likelihood function in Cases I and II greatly helps in preparing our intuition and solving the more complex 3-D optimization problem. Therefore, we next proceed with a stepwise approach by considering these four cases separately one-by-one.

a. Case I: τ known, ϕ known

Without losing any generalization, the likelihood function in this case can be obtained by making $\phi = 0$ in (3.23). From the form of the likelihood function, we can see that it is nonzero only over a certain support region defined by the limits of the indicator function $I[.]$. Since τ is fixed and known, the set of constraints in (3.23), namely

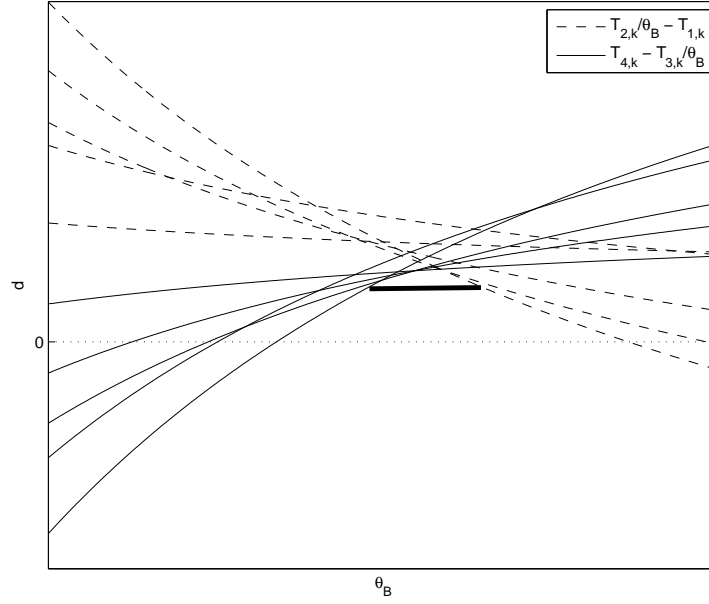


Fig. 9. Nonzero likelihood region shown as the solid line.

$\tau > 0$, $\omega > 0$ and

$$\tau \leq \frac{T_{2,k}}{\omega} - T_{1,k}, \quad k = 1, \dots, N, \quad (3.25)$$

$$\tau \leq T_{4,k} - \frac{T_{3,k}}{\omega}, \quad k = 1, \dots, N, \quad (3.26)$$

can be equivalently put in the form:

$$\begin{aligned} \tau &> 0, \quad \omega > 0, \\ \frac{T_{3,k}}{T_{4,k} - \tau} &\leq \omega \leq \frac{T_{2,k}}{T_{1,k} + \tau}, \quad k = 1, \dots, N. \end{aligned} \quad (3.27)$$

Fig. 9 shows various upper-bounds (3.25) and (3.26) of the likelihood support region in the plane (τ, ω) , and the solid line is the region over which the likelihood function has to be maximized. It is evident from the figure that for a known fixed τ , the likelihood function depends on the unknown ω only and is maximized by taking ω

Algorithm 1 Finding $\hat{\omega}$ for τ known, ϕ known

- 1: Find $\omega_k = \frac{T_{3,k}}{T_{4,k} - \tau}$, for $k = 1, \dots, N$;
 - 2: $j = \arg \max_k \{\omega_k\}$;
 - 3: $\hat{\omega} = \hat{\omega}_j$;
-

as small as possible. This is because the factor $\sum_{k=1}^N (T_{2,k} - T_{3,k})$ in (3.23) is always negative. Therefore, the smallest value of ω over the solid line, as shown in Fig. 9, is the MLE $\hat{\omega}$, which coincides with one of the curves $\tau = T_{4,k} - T_{3,k}/\omega$, $k = 1, \dots, N$. Let j ($1 \leq j \leq N$) denote the index of the curve on which the MLE is achieved. Thus, from (3.27), $j = \arg \max_k \{T_{3,k}/(T_{4,k} - \tau)\}$ and

$$\hat{\omega} = \frac{T_{3,j}}{T_{4,j} - \tau}.$$

The index j , which gives the set of timestamps $\{T_{3,j}, T_{4,j}\}$ required for finding the MLE, is the one which gives the minimum possible $\hat{\omega}$ over the allowable region. Since τ is known, we can find j , and hence the corresponding $\hat{\omega}$, by Algorithm 1.

Algorithm 1 utilizes the fact that the solid line cuts all the curves $T_{4,k} - T_{3,k}/\omega$, $k = 1, \dots, N$ but the likelihood function is zero beyond its intersection with the first curve, which is the maximum of these intersections and therefore gives the MLE. Note that in doing so, a total number of N values need to be compared. To simplify the exposition, in what follows we will use the terminology *the curves* $T_{4,k} - T_{3,k}/\omega$, $k = 1, \dots, N$, instead of *the curves* $\tau = T_{4,k} - T_{3,k}/\omega$, $k = 1, \dots, N$.

b. Case II: τ unknown, ϕ known

The likelihood function in this case is similar to Case I, but with one major difference: the fixed delay τ is unknown. The shaded region in Fig. 10 is the subset of Θ over which the likelihood function is nonzero. It can be described in terms of the following

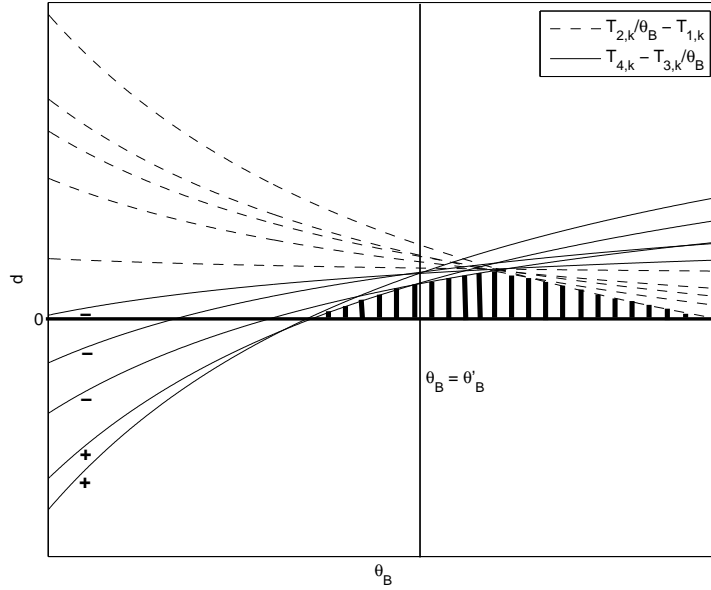


Fig. 10. Support region of the likelihood function shown as the shaded area. Also shown to the left in this figure is the sign of the term $2NT_{3,j} - \sum_{k=1}^N (T_{3,k} - T_{2,k})$ for each $j = 1, \dots, N$.

constraints:

$$\begin{aligned} \tau &> 0, \quad \omega > 0, \\ \tau &\leq \frac{T_{2,k}}{\omega} - T_{1,k}, \quad k = 1, \dots, N, \end{aligned} \quad (3.28)$$

$$\tau \leq T_{4,k} - \frac{T_{3,k}}{\omega}, \quad k = 1, \dots, N. \quad (3.29)$$

This likelihood function in (3.23) is maximized by making its argument:

$$\xi = \sum_{k=1}^N \frac{T_{2,k} - T_{3,k}}{\omega} - \sum_{k=1}^N (T_{1,k} - T_{4,k}) - 2Nd, \quad (3.30)$$

as small as possible. Although Fig. 10 shows only the support region and not the

likelihood function itself, ξ can be linked to this figure by rewriting it in the form

$$\xi = \sum_{k=1}^N \left(\frac{T_{2,k}}{\omega} - T_{1,k} \right) + \sum_{k=1}^N \left(T_{4,k} - \frac{T_{3,k}}{\omega} \right) - 2Nd,$$

and noting that for any $\omega = \omega'$, ξ is the sum of the ordinates of all points on the curves $(T_{2,k}/\omega - T_{1,k})$, $k = 1, \dots, N$, and $(T_{4,k} - T_{3,k}/\omega)$, $k = 1, \dots, N$, intercepting the vertical line $\omega = \omega'$, minus $2N$ times $\hat{\tau}$ (which is the intersection of $\omega = \omega'$ with either $\min_{1 \leq k \leq N} (T_{2,k}/\omega - T_{1,k})$ or $\min_{1 \leq k \leq N} (T_{4,k} - T_{3,k}/\omega)$ as proved in Lemma 1 below). Utilizing the fact that ξ depends on two parameters, ω and τ , we will now derive the MLE with the help of the following four lemmas:

Lemma 1. *The MLE $\hat{\tau}$ lies on either $\min_{1 \leq k \leq N} (T_{2,k}/\omega - T_{1,k})$ or $\min_{1 \leq k \leq N} (T_{4,k} - T_{3,k}/\omega)$, i.e., on the boundary of the support region.*

Proof. This can be proved by contradiction. Let us assume that the $\hat{\tau}$ does not lie on the boundary, but somewhere else inside the support region. Then for some minimizing $\hat{\omega}$, ξ can be further decreased by increasing $\hat{\tau}$ to the top of the allowable region (which coincides with one of the the above mentioned curves) for the same $\hat{\omega}$, hence a contradiction. \square

Lemma 2. *The MLE $\hat{\tau}$ lies either on the uppermost vertex formed by the intersection of the curves $\min_{1 \leq k \leq N} (T_{2,k}/\omega - T_{1,k})$ and $\min_{1 \leq k \leq N} (T_{4,k} - T_{3,k}/\omega)$ (shown as point A in Fig. 11) or on one of the vertices formed by the intersection of the curves $(T_{4,k} - T_{3,k}/\omega)$, $k = 1, \dots, N$ (shown as points B, C, etc. in Fig. 11).*

Proof. It is straightforward to notice from (3.30) that when $T_{2,k} = T_{3,k}$, for all k , ξ can be minimized by making τ as large as possible, which is the intersection of the curves $\min_{1 \leq k \leq N} (T_{2,k}/\omega - T_{1,k})$ and $\min_{1 \leq k \leq N} (T_{4,k} - T_{3,k}/\omega)$. Hence, the MLE $(\hat{\tau}, \hat{\omega})$ is

$$\hat{\tau} = \frac{T_{2,i}T_{4,j} - T_{1,i}T_{3,j}}{T_{2,i} + T_{3,j}}, \quad (3.31)$$

and

$$\hat{\omega} = \frac{T_{2,i} + T_{3,j}}{T_{1,i} + T_{4,j}}, \quad (3.32)$$

where the i, j represent the indices of $\min_{1 \leq k \leq N} (T_{2,k}/\omega - T_{1,k})$ and $\min_{1 \leq k \leq N} (T_{4,k} - T_{3,k}/\omega)$, respectively, intersecting at the maximum $\hat{\tau}$ (which is the uppermost vertex shown as point A in Fig. 11). Note that in order to find this MLE, a total number of N^2 intersections have to be compared.

When $T_{2,k} \neq T_{3,k}$, for some k , the problem becomes a little involved. From Lemma 1, we know that $\hat{\tau}$ lies somewhere on the boundary of the support region. Notice further that according to (3.30) in order to minimize ξ it is necessary to select τ as large as possible and ω as small as possible.

Suppose that $\hat{\tau}$ lies on $\min_{1 \leq k \leq N} (T_{2,k}/\omega - T_{1,k})$ and let $i = \arg \min_{1 \leq k \leq N} (T_{2,k}/\omega - T_{1,k})$ corresponding to the maximum τ (i.e., point A in Fig. 11), then from (3.30) ξ can be written as

$$\begin{aligned} \xi &= \sum_{k=1}^N \frac{T_{2,k} - T_{3,k}}{\omega} - \sum_{k=1}^N (T_{1,k} - T_{4,k}) - 2N \left(\frac{T_{2,i}}{\omega} - T_{1,i} \right), \\ &= \frac{1}{\omega} \sum_{k=1}^N (T_{2,k} - T_{3,k} - 2T_{2,i}) - \sum_{k=1}^N (T_{1,k} - T_{4,k} - 2T_{1,i}). \end{aligned} \quad (3.33)$$

Since the term $\sum_{k=1}^N (T_{2,k} - T_{3,k} - 2T_{2,i})$ is always negative, ξ can be minimized by taking ω as small as possible on $\min_{1 \leq k \leq N} (T_{2,k}/\omega - T_{1,k})$. Hence, $\hat{\tau}$ and $\hat{\omega}$ in this general case are *equal to* or *less than* the MLE given by (3.31) and (3.32), respectively (i.e., either on point A shown in Fig. 11 or to the left of it). An alternative justification for the fact that $\hat{\tau}$ and $\hat{\omega}$ are *equal to* or *less than* the MLE given by (3.31) and (3.32), respectively, is to assume by contradiction that $\hat{\tau}$ lies on $\min_{1 \leq k \leq N} (T_{2,k}/\omega - T_{1,k})$, with $i = \arg \min_{1 \leq k \leq N} (T_{2,k}/\omega - T_{1,k})$, which does not correspond to the maximum τ (i.e., not on the curve passing through point A in Fig. 11). According to (3.33), ξ is minimized by

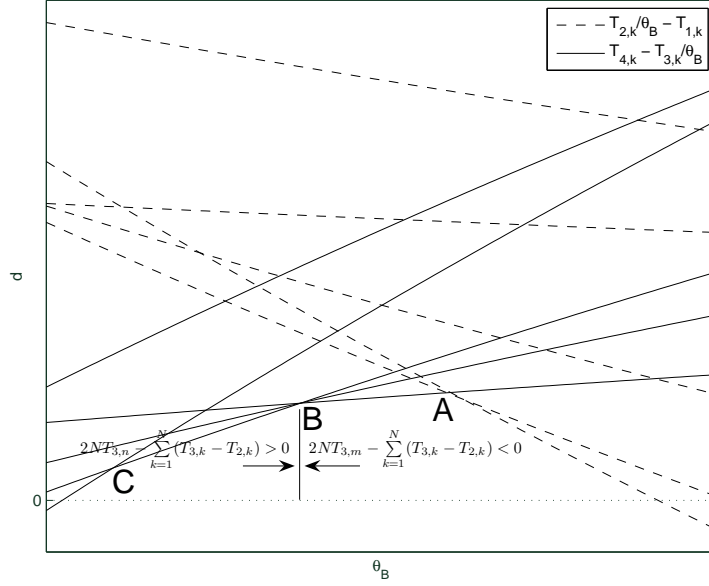


Fig. 11. Zoomed in version of the support region of the likelihood function.

choosing ω as small as possible. Taking into account the continuity of ξ with respect to ω and τ , one can show that ξ is monotonically decreasing as long as ω is decreased until it reaches the value corresponding to the point A.

Now suppose that $\hat{\tau}$ lies on $\min_{1 \leq k \leq N} (T_{4,k} - T_{3,k}/\omega)$ and let $j = \arg \min_{1 \leq k \leq N} (T_{4,k} - T_{3,k}/\omega)$ corresponding to the maximum τ (i.e., point A in Fig. 11), then ξ can be written as

$$\begin{aligned} \xi &= \sum_{k=1}^N \frac{T_{2,k} - T_{3,k}}{\omega} - \sum_{k=1}^N (T_{1,k} - T_{4,k}) - 2N \left(T_{4,j} - \frac{T_{3,j}}{\omega} \right), \\ &= \frac{1}{\omega} \sum_{k=1}^N (T_{2,k} - T_{3,k} + 2T_{3,j}) - \sum_{k=1}^N (T_{1,k} - T_{4,k} + 2T_{4,j}). \end{aligned} \quad (3.34)$$

From (3.34), it is clear that ξ can be minimized by taking the largest possible ω if $\sum_{k=1}^N (T_{2,k} - T_{3,k} + 2T_{3,j})$ is positive and by taking the smallest possible ω if $\sum_{k=1}^N (T_{2,k} - T_{3,k} + 2T_{3,j})$

is negative as depicted by Fig. 11. Hence, for

$$2NT_{3,j} > \sum_{k=1}^N (T_{3,k} - T_{2,k}),$$

MLE is again given by (3.31) and (3.32). And for

$$2NT_{3,j} < \sum_{k=1}^N (T_{3,k} - T_{2,k}),$$

MLE is given by the intersection of the curves $(T_{4,m} - T_{3,m}/\omega)$ and $(T_{4,n} - T_{3,n}/\omega)$ (denoting the intersections of the curves $\tau = T_{4,k} - T_{3,k}/\omega$ and $\tau = T_{4,l} - T_{3,l}/\omega$ as $\tau^{k,l}$, and $\tau^{k,l}$ satisfy the constraints (3.28) and (3.29)), where

$$(m, n) = \arg \max_{k,l} \{ \tau^{k,l} \mid 2NT_{3,k} < \sum_{r=1}^N (T_{3,r} - T_{2,r}); 2NT_{3,l} > \sum_{r=1}^N (T_{3,r} - T_{2,r}) \}. \quad (3.35)$$

Basically, the indices (m, n) in (3.35) identify the first vertex of the support region located to the left of the vertex A for which a change of sign occurs in $2NT_{3,n} - \sum_{r=1}^N (T_{3,r} - T_{2,r})$. In Fig. 4, this vertex is represented by the point B, and the MLE $(\hat{\tau}, \hat{\omega})$ in this case is given by

$$\hat{\tau} = T_{4,m} - \frac{T_{3,m}(T_{4,m} - T_{4,n})}{T_{3,m} - T_{3,n}}, \quad (3.36)$$

and

$$\hat{\omega} = \frac{T_{3,m} - T_{3,n}}{T_{4,m} - T_{4,n}}, \quad (3.37)$$

□

Lemma 3. *To the left of the point where $\min_{1 \leq k \leq N} (T_{2,k}/\omega - T_{1,k})$ and $\min_{1 \leq k \leq N} (T_{4,k} - T_{3,k}/\omega)$ intersect (i.e., point A shown in Fig. 11), the boundary of the support region is formed by the curves $(T_{4,k} - T_{3,k}/\omega)$, $k = 1, \dots, N$ in such a way that as ω increases, a curve $(T_{4,m} - T_{3,m}/\omega)$ forms the new boundary of the support region after intersecting the*

curve $(T_{4,n} - T_{3,n}/\omega)$ if and only if $m < n$.

Proof. The curve $(T_{4,N} - T_{3,N}/\omega)$ starts as the most negative for small ω and ends up as the largest positive asymptotically approaching $T_{4,N}$ as ω increases. Similarly, the curve $(T_{4,1} - T_{3,1}/\omega)$ starts as the least negative for small ω and ends up as the smallest positive asymptotically approaching $T_{4,1}$ as ω increases. All the curves $(T_{4,k} - T_{3,k}/\omega)$ $k = 1, \dots, N$, are arranged in descending order for small ω and in ascending order for large ω and they intersect each other somewhere around the true value of ω . Since the slope of each curve $(T_{4,k} - T_{3,k}/\omega)$, $k = 1, \dots, N$ is $T_{3,k}/\omega^2$, the slope of the curve with index m is lesser than the slope of the curve with index n if $m < n$. Therefore, as ω increases, a curve can form the new boundary of the support region by intersecting another curve only if its index is lower than the previous one. \square

Lemma 4. *The MLE $(\hat{\tau}, \hat{\omega})$, whether (3.31) and (3.32) or (3.36) and (3.37), is unique.*

Proof. Note that the likelihood function is continuous on the boundary of the support region because different curves intersect each other on the vertices due to which there will be no jumps in ξ and subsequently in the likelihood function. Now considering the fact that $2NT_{3,j} > \sum_{k=1}^N (T_{3,k} - T_{2,k})$ for $j = N$, let

$$q = \arg \max_j \{T_{3,j} \mid 2NT_{3,j} < \sum_{k=1}^N (T_{3,k} - T_{2,k})\}.$$

Then it must also be true that $2NT_{3,j} < \sum_{k=1}^N (T_{3,k} - T_{2,k}) \forall j < q$, i.e., for $j = 1, \dots, q-1$ and $2NT_{3,j} > \sum_{k=1}^N (T_{3,k} - T_{2,k}) \forall j > q$, i.e., for $j = q+1, \dots, N$. Fig. 10 shows the sign of the term $2NT_{3,j} - \sum_{k=1}^N (T_{3,k} - T_{2,k})$ for each $j = 1, \dots, N$. There

Algorithm 2 Finding $\hat{\omega}$ and $\hat{\tau}$ for τ unknown, ϕ known

- 1: Find $\tau^{k,l} = \frac{T_{2,k}T_{4,l} - T_{1,k}T_{3,l}}{T_{2,k} + T_{3,l}}$; $\omega^{k,l} = \frac{T_{2,k} + T_{3,l}}{T_{1,k} + T_{4,l}}$; $\forall k = 1, \dots, N$ and $\forall l = 1, \dots, N$.
 - 2: $(i, j) = \arg \min_{k,l} \{\tau^{k,l}\}$;
 - 3: **if** $2NT_{3,j} > \sum_{k=1}^N (T_{3,k} - T_{2,k})$ **then**
 - 4: $\hat{\tau} = \tau^{i,j}$; $\hat{\omega} = \omega^{i,j}$;
 - 5: **else**
 - 6: $k = j$;
 - LABEL:
 - 7: Find $\tau^{k,l} = T_{4,k} - \frac{T_{3,k}(T_{4,k} - T_{4,l})}{T_{3,k} - T_{3,l}}$; $\omega^{k,l} = \frac{T_{3,k} - T_{3,l}}{T_{4,k} - T_{4,l}}$; $\forall l = 1, \dots, N$.
 - 8: $m = k$; $n = \arg \max_l \{\omega^{k,l} \mid \omega^{k,l} < \omega^{i,j}\}$;
 - 9: **if** $2NT_{3,n} > \sum_{k=1}^N (T_{3,k} - T_{2,k})$ **then**
 - 10: $\hat{\tau} = \tau^{m,n}$; $\hat{\omega} = \omega^{m,n}$;
 - 11: **else**
 - 12: $k = n$;
 - 13: goto LABEL;
 - 14: **end if**
 - 15: **end if**
-

will always be just one change, if any, in the sign of this term from positive to negative. Therefore, ξ can be minimized by making ω as large as possible on $(T_{4,q+1} - T_{3,q+1}/\omega)$ and as small as possible on the curve $(T_{4,q} - T_{3,q}/\omega)$ (or on $(T_{2,i}/\omega - T_{1,i})$ if there is no such q) as shown in Fig. 10.

This fact, combined with Lemma 3, proves that the intersection of the curves forming the MLE is always unique. \square

It should be noted that under the most likely scenario, when *Node B* is sending

its timestamps to *Node A* after short delays, MLE will be given by (3.31) and (3.32), but in the usually unlikely scenario of *Node B* waiting a long period of time before sending one of its timestamps to *Node A*, (3.36) and (3.37) can be the MLE only if $2NT_{3,j} < \sum_{k=1}^N (T_{3,k} - T_{2,k})$. Note that in this case, in addition to previous N^2 intersections, $N - 1$ more intersections have to be compared for each j satisfying $2NT_{3,j} < \sum_{k=1}^N (T_{3,k} - T_{2,k})$. The whole procedure for finding this MLE is summarized in Algorithm 2. This algorithm proceeds in precisely the same steps as described above.

Now that we have obtained some insight into this problem for ϕ known, we next proceed with the situation when ϕ is unknown.

c. Case III: τ known, ϕ unknown

The likelihood function in this case is the same as (3.23), where τ is fixed and known. The region over which the likelihood function is nonzero is given by indicator function $I[\cdot]$ in (3.23) and shown in Fig. 12. This 3-D support region is dramatically more complex than what we observed in the first two cases. It is also evident from (3.23) that ξ is the same as in previous cases and the likelihood function can again be maximized by minimizing ξ . Since $\sum_{k=1}^N (T_{2,k} - T_{3,k})$ is always negative and τ is given, ξ can be minimized by taking ω as small as possible. To find this minimum ω , we take a horizontal slice from this 3-D support region at the constant τ . This gives an aerial view of the 2-D region shown in Fig. 13 highlighting the relation between ω versus ϕ for the known τ . Therefore, in accordance with (3.23), we can express the support of the likelihood function in the form of the following constraints:

$$-\infty < \phi < \infty, \quad (3.38)$$

$$\frac{T_{3,k} - \phi}{T_{4,k} - \tau} \leq \omega \leq \frac{T_{2,k} - \phi}{T_{1,k} + \tau}; \quad k = 1, \dots, N. \quad (3.39)$$

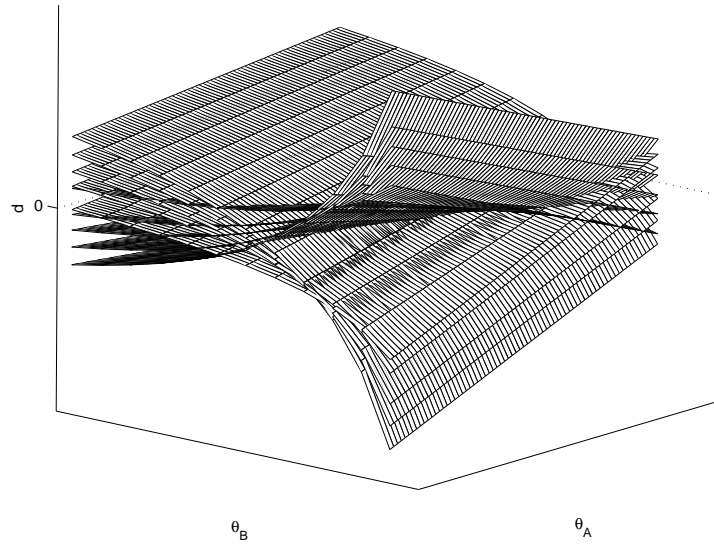


Fig. 12. τ as a function of ϕ and ω .

These constraints can be viewed as ω being a monotonically decreasing function of ϕ $\forall k$ due to the positivity of $(T_{1,k} + \tau)$ and $(T_{4,k} - \tau)$, and the shaded region is the one which satisfies these constraints.

Lemma 5. *Of all the intersections of $(T_{2,k} - \phi)/(T_{1,k} + \tau)$ with $(T_{3,k} - \phi)/(T_{4,k} - \tau)$, only two points satisfy the constraints (3.38) and (3.39) in a way that they represent the starting and ending points of the support region and the point with minimum ω is the one with maximum ϕ .*

Proof. Consider the curves $(T_{2,k} - \psi)/(T_{1,k} + \tau)$ with $(T_{3,k} - \psi)/(T_{4,k} - \tau)$ as a function of ψ in order to avoid confusion between the actual unknown parameter ϕ and the variable with respect to which the above functions are drawn. Now utilizing (3.4)

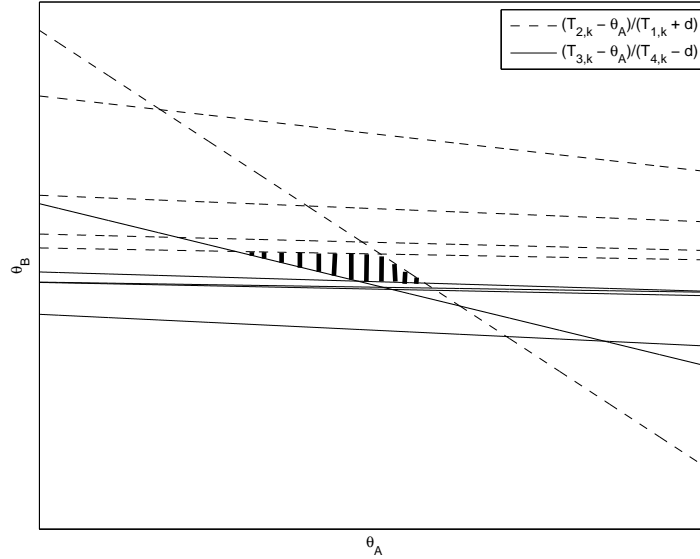


Fig. 13. ω as a function of ϕ for constant τ .

and (3.5), we can write

$$\begin{aligned} \frac{T_{2,k} - \psi}{T_{1,k} + \tau} &= \left(1 + \frac{X_k}{T_{1,k} + \tau}\right) \omega + \frac{\phi - \psi}{T_{1,k} + \tau}, & k = 1, \dots, N, \\ \frac{T_{3,k} - \psi}{T_{4,k} - \tau} &= \left(1 - \frac{Y_k}{T_{4,k} - \tau}\right) \omega + \frac{\phi - \psi}{T_{4,k} - \tau}, & k = 1, \dots, N. \end{aligned}$$

It is clear that when $\psi = \phi$, $(T_{2,k} - \psi)/(T_{1,k} + \tau) > (T_{3,k} - \psi)/(T_{4,k} - \tau) \forall k$. Therefore, a support region *does* exist where the constraints (3.38) and (3.39) are satisfied. Now the slopes and y-intercepts of the straight lines $(T_{2,k} - \psi)/(T_{1,k} + \tau)$ are $-(T_{1,k} + \tau)^{-1}$ and $[1 + X_k(T_{1,k} + \tau)^{-1}]\omega + \phi(T_{1,k} + \tau)^{-1}$ respectively, and the slopes and y-intercepts of the straight lines $(T_{3,k} - \psi)/(T_{4,k} - \tau)$ are $-(T_{4,k} - \tau)^{-1}$ and $[1 - Y_k(T_{4,k} - \tau)^{-1}]\omega + \phi(T_{4,k} - \tau)^{-1}$ respectively. The y-intercepts can attain any value depending on the random delays X_k , Y_k and the sign and magnitude of ϕ , but there is a set pattern in the slopes of these lines. According to the model (see

Fig. 5), it is always true that

$$(T_{1,1} + \tau)^{-1} > (T_{4,1} - \tau)^{-1} > (T_{1,2} + \tau)^{-1} > \dots > \\ (T_{4,N-1} - \tau)^{-1} > (T_{1,N} + \tau)^{-1} > (T_{4,N} - \tau)^{-1}.$$

This is because $T_{1,1} < T_{4,1} < T_{1,2} < \dots < T_{4,N-1} < T_{1,N} < T_{4,N}$. Due to the alternating slopes, the lines $(T_{2,k} - \psi)/(T_{1,k} + \tau)$ and $(T_{3,k} - \psi)/(T_{4,k} - \tau)$ for every k intersect each other on at least one point. According to the order of the slopes, both to the left and right of $\psi = \phi$, the support region ends after the first intersection. Therefore, there are exactly two points, $(\phi', \omega')_1$ and $(\phi', \omega')_2$, which define the starting and ending point of the support region. In addition, the point corresponding to minimum ω' is the one with maximum ϕ' since all the straight lines always have negative slopes. \square

We can minimize ξ by taking the intersection of $\min_{1 \leq k \leq N} (T_{2,k} - \phi)/(T_{1,k} + \tau)$ and $\max_{1 \leq l \leq N} (T_{3,l} - \phi)/(T_{4,l} - \tau)$ at minimum possible $\hat{\omega}$, which gives the MLE $(\hat{\phi}, \hat{\omega})$ as

$$\hat{\phi} = \frac{T_{3,j}(T_{1,i} + \tau) - T_{2,i}(T_{4,j} - \tau)}{(T_{1,i} + \tau) - (T_{4,j} - \tau)}, \\ \hat{\omega} = \frac{T_{2,i} - T_{3,j}}{(T_{1,i} + \tau) - (T_{4,j} - \tau)},$$

where the indices (i, j) are the ones whose intersection gives the minimum allowed $\hat{\omega}$. Algorithm 3 presents in detail the steps that are required for finding this MLE.

Algorithm 3 first finds all the intersections and chooses two candidate points $(\phi', \omega')_1$ and $(\phi', \omega')_2$ such that $\omega' \leq (T_{2,k} - \phi)/(T_{1,k} + \tau) \forall k$ and $\omega' \geq (T_{3,l} - \phi)/(T_{4,l} - \tau) \forall l$. These are the starting and ending points of the nonzero likelihood region as proved in Lemma 5 above and the point with minimum ω (which corresponds to the one with maximum ϕ) is chosen.

Algorithm 3 Finding $\hat{\phi}$ and $\hat{\omega}$ for τ known, ϕ unknown

- 1: Find $\phi^{k,l} = \frac{T_{3,l}(T_{1,k}+\tau)-T_{2,k}(T_{4,l}-\tau)}{(T_{1,k}+\tau)-(T_{4,l}-\tau)}$; $\omega^{k,l} = \frac{T_{2,k}-T_{3,l}}{(T_{1,k}+\tau)-(T_{4,l}-\tau)}$; $\forall k = 1, \dots, N$ and $\forall l = 1, \dots, N$;
 - 2: $(i, j) = \{(k, l) \mid \omega^{k,l} \leq \frac{T_{2,r}-\phi^{k,l}}{T_{1,r}+\tau} \forall r \text{ and } \omega^{k,l} \geq \frac{T_{3,r}-\phi^{k,l}}{T_{4,r}-\tau} \forall r\}$;
 - 3: $(m, n) = \{(k, l) \mid (k, l) \neq (i, j), \omega^{k,l} \leq \frac{T_{2,r}-\phi^{k,l}}{T_{1,r}+\tau} \forall r \text{ and } \omega^{k,l} \geq \frac{T_{3,r}-\phi^{k,l}}{T_{4,r}-\tau} \forall r\}$;
 - 4: $\hat{\omega} = \min\{\omega^{i,j}, \omega^{m,n}\}$; $\hat{\phi} = \max\{\phi^{i,j}, \phi^{m,n}\}$;
-

d. Case IV: τ unknown, ϕ unknown

In this case, all of τ, ϕ and ω are unknown and have to be jointly estimated. The likelihood function in this case is the same as in (3.23) but τ is unknown. The region where the likelihood function is nonzero can be expressed in the form of the following constraints:

$$-\infty < \phi < \infty, \tau > 0, \omega > 0,$$

$$\tau \leq \frac{T_{2,k} - \phi}{\omega} - T_{1,k}, \quad k = 1, \dots, N, \quad (3.40)$$

$$\tau \leq T_{4,k} - \frac{T_{3,k} - \phi}{\omega}, \quad k = 1, \dots, N. \quad (3.41)$$

Within the constraint $\tau > 0$, $(T_{2,k} - \phi)/\omega - T_{1,k}$ are monotonically decreasing functions of ϕ and $\omega \forall k$, and $T_{4,k} - (T_{3,k} - \phi)/\omega$ are monotonically increasing functions of ϕ and $\omega \forall k$ as shown in Fig. 12. It is clear from the same figure that the nonzero likelihood region is similar in shape to a dome if we look at it standing on (ϕ, ω) plane. Lemma 1 asserts that the MLE $(\hat{\tau}, \hat{\phi}, \hat{\omega})$ should lie somewhere on the ceiling of this dome. The lines on (ϕ, ω) plane, on which the intersections of the surfaces lie are given by

$$\phi = \frac{1}{2} [(T_{2,k} + T_{3,l}) - \omega(T_{1,k} + T_{4,l})], \quad k = 1, \dots, N; \quad l = 1, \dots, N, \quad (3.42)$$

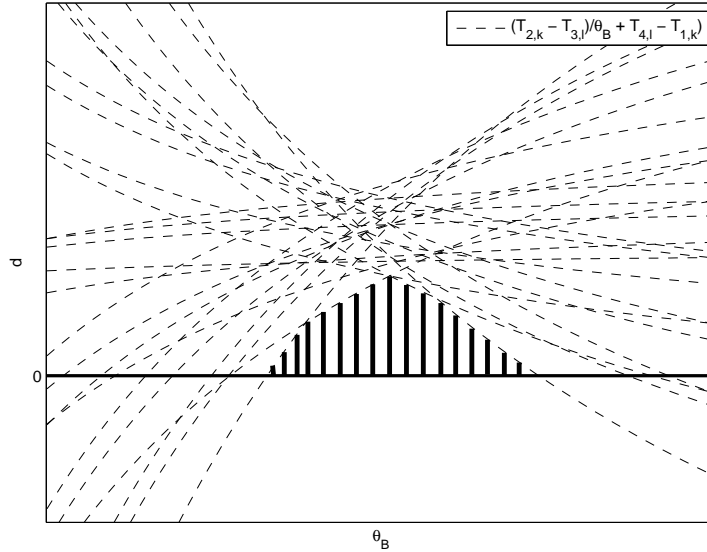


Fig. 14. τ as a function of ω only.

or equivalently

$$\omega = \frac{T_{2,k} + T_{3,l} - 2\phi}{T_{1,k} + T_{4,l}}, \quad k = 1, \dots, N; l = 1, \dots, N. \quad (3.43)$$

Note that putting $\omega = 1$ (the case when there is no clock skew) in (3.42) and taking the minimum results in the MLE $\hat{\phi}$ in (3.24) derived by [18]. Although τ is a function of both ϕ and ω , it can be written as a function of either ϕ only or ω only by utilizing this linear relationship between these two parameters. Fig. 14 shows the *imaginary* 2-D region where τ is drawn as a function of ω only and Fig. 15 shows the *imaginary* 2-D region where τ is drawn as a function of ϕ only. Note that these are actually 3-D plots, but the points on the bottom two axes (ϕ, ω) are replaced with $(\frac{1}{2} [(T_{2,k} + T_{3,l}) - \omega(T_{1,k} + T_{4,l})], \omega)$ and $(\phi, (T_{2,k} + T_{3,l} - 2\phi)/(T_{1,k} + T_{4,l}))$ in Fig. 14 and Fig. 15, respectively.

Over the line (3.42), τ is given by

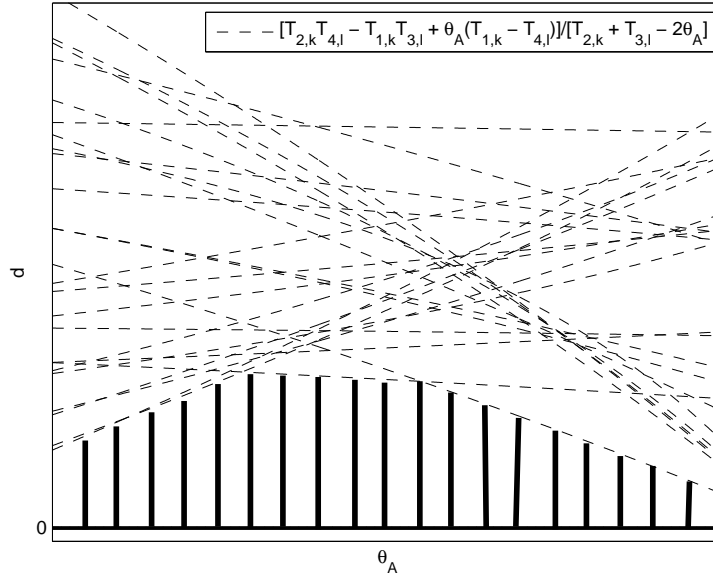


Fig. 15. τ as a function of ϕ only.

$$\tau \leq \frac{1}{2} \left[\frac{T_{2,k} - T_{3,l}}{\omega} + (T_{4,l} - T_{1,k}) \right], \quad k = 1, \dots, N; \quad l = 1, \dots, N. \quad (3.44)$$

Note that putting $\omega = 1$ and taking the min results in MLE $\hat{\tau}$ given by [18]. And over the line (3.43), τ is given by

$$\tau \leq \frac{T_{2,k}T_{4,l} - T_{1,k}T_{3,l} + \phi(T_{1,k} - T_{4,l})}{T_{2,k} + T_{3,l} - 2\phi}, \quad k = 1, \dots, N; \quad l = 1, \dots, N. \quad (3.45)$$

A closer look at (3.45) reveals that its RHS goes to $-\infty$ or $+\infty$ respectively at $\phi = (T_{2,i} + T_{3,j})/2$ according to the negative or positive sign of the numerator. But the constraint $\tau > 0$ automatically restricts the nonzero likelihood region well before even the first discontinuity of this kind as shown in Fig. 15.

Estimating τ and ω : Consider the set of N^2 curves given in (3.44) and plotted in Fig. 14. Since the signs of $T_{2,k} - T_{3,l}$ and $T_{4,l} - T_{1,k}$ are always opposite, $N(N - 1)/2$ of these curves have positive numerator in the term involving ω

and negative constant term, while the remaining $N(N + 1)/2$ have negative numerator in the term involving ω and positive constant term. Based on this observation, (3.44) can be written in the form of two sets of inequalities such that $(T_{2,k} - T_{3,l}) > 0$ for one set and $(T_{2,k} - T_{3,l}) < 0$ for the other as shown in Fig. 14. Then the current scenario assumes quite a similar form to the set of constraints (3.28) and (3.29). Therefore, initially a total of $[N(N - 1)/2][N(N + 1)/2] = N^2(N^2 - 1)/4$ intersections (denoted by $\tau^{k,l,m,n}$ in Algorithm 4) are to be compared. Lemmas 1, 2, 3 and 4 are then similarly true for these sets of inequalities and the MLEs can be derived by following a similar procedure. Let us denote $\{\min_{1 \leq k, l \leq N} \frac{1}{2} [(T_{2,k} - T_{3,l})/\omega + (T_{4,l} - T_{1,k})] \mid (T_{2,k} - T_{3,l}) > 0\}$ as $(T_{2,i} - T_{3,j})/2\omega + (T_{4,j} - T_{1,i})/2$ and $\{\min_{1 \leq k, l \leq N} \frac{1}{2} [(T_{2,k} - T_{3,l})/\omega + (T_{4,l} - T_{1,k})] \mid (T_{2,k} - T_{3,l}) < 0\}$ as $(T_{2,m} - T_{3,n})/2\omega + (T_{4,n} - T_{1,m})/2$. Then if $\sum_{k=1}^N [T_{2,k} - T_{3,k} - (T_{2,m} - T_{3,n})]$ is positive, the MLE $(\hat{\tau}, \hat{\omega})$ is the intersection of this curve with the one discussed above, i.e.,

$$\hat{\tau} = \frac{1}{2} \left[\frac{(T_{2,i} - T_{3,j})(T_{1,i} - T_{1,m}) + (T_{4,n} - T_{4,j})}{(T_{2,i} - T_{2,m}) + (T_{3,n} - T_{3,j})} + (T_{4,j} - T_{1,i}) \right],$$

and

$$\hat{\omega} = \frac{(T_{2,i} - T_{2,m}) + (T_{3,n} - T_{3,j})}{(T_{1,i} - T_{1,m}) + (T_{4,n} - T_{4,j})}, \quad (3.46)$$

Otherwise, if $\sum_{k=1}^N [T_{2,k} - T_{3,k} - (T_{2,m} - T_{3,n})]$ is negative, then the MLE is the intersection of the curves $(T_{2,p} - T_{3,q})/2\omega + (T_{4,q} - T_{1,p})/2$ and $(T_{2,r} - T_{3,s})/2\omega + (T_{4,s} - T_{1,r})/2$ (denoting the intersections of the curves in (3.44) as $\tau^{k,l,m,n}$, $\forall (k, l, m, n)$, and $\tau^{k,l,m,n}$ satisfy the constraints (3.40) and (3.41)), where

$$(p, q, r, s) = \arg \max_{k,l,m,n} \{ \tau^{k,l,m,n} \mid N(T_{3,l} - T_{2,k}) < \sum_{k=1}^N (T_{3,k} - T_{2,k}); \\ N(T_{3,n} - T_{2,m}) > \sum_{k=1}^N (T_{3,k} - T_{2,k}) \}.$$

Algorithm 4 Finding $\hat{\phi}$, $\hat{\omega}$ and $\hat{\tau}$ for τ unknown, ϕ unknown

1: $(m, n) = (1, N)$;

LABEL:

2: Find $\tau^{k,l,m,n} = \frac{1}{2} \left[\frac{(T_{2,k}-T_{3,l})((T_{1,k}-T_{1,m})+(T_{4,n}-T_{4,l}))}{(T_{2,k}-T_{2,m})+(T_{3,n}-T_{3,l})} + (T_{4,l} - T_{1,k}) \right]$; $\omega^{k,l,m,n} = \frac{(T_{2,k}-T_{2,m})+(T_{3,n}-T_{3,l})}{(T_{1,k}-T_{1,m})+(T_{4,n}-T_{4,l})}$, $\forall (k, l) \neq (m, n)$;

3: $(p, q) = \arg \min_{k,l} \{\tau^{k,l,m,n}\}$

4: **if** $T_{2,p} - T_{3,q} > 0$ **then**

5: $\hat{\tau} = \tau^{p,q,m,n}$; $\hat{\omega} = \omega^{p,q,m,n}$; $\hat{\phi} = \frac{1}{2} [(T_{2,p} + T_{3,q}) - \hat{\omega}(T_{1,p} + T_{4,q})]$;

6: **else**

7: **if** $N(T_{2,p} - T_{3,q}) > \sum_{k=1}^N (T_{3,k} - T_{2,k})$ **then**

8: $\hat{\tau} = \tau^{p,q,m,n}$; $\hat{\omega} = \omega^{p,q,m,n}$; $\hat{\phi} = \frac{1}{2} [(T_{2,p} + T_{3,q}) - \hat{\omega}(T_{1,p} + T_{4,q})]$;

9: **else**

10: Remove (m, n) curve;

11: $(m, n) = (p, q)$;

12: goto LABEL;

13: **end if**

14: **end if**

Hence, here the MLE $(\hat{\tau}, \hat{\omega})$ is

$$\hat{\tau} = \frac{1}{2} \left[\frac{(T_{2,p} - T_{3,q})((T_{1,p} - T_{1,r}) + (T_{4,s} - T_{4,q}))}{(T_{2,p} - T_{2,r}) + (T_{3,s} - T_{3,q})} + (T_{4,q} - T_{1,p}) \right],$$

and

$$\hat{\omega} = \frac{(T_{2,p} - T_{2,r}) + (T_{3,s} - T_{3,q})}{(T_{1,p} - T_{1,r}) + (T_{4,s} - T_{4,q})}, \quad (3.47)$$

The complete procedure for finding the MLE is described in Algorithm 4. Although a modified Algorithm 2 can be used in this case, we present this alternative

algorithm for the sake of completion. It starts from the curve for which $(T_{2,m} - T_{3,n})$ is minimum, i.e., $(T_{2,1} - T_{3,N})$ and then compares its intersections with other curves. It keeps on replacing this curve with the one giving the next minimum $\tau^{k,l,m,n}$ within the constraints until the MLE is found according to the procedure described before.

Estimating ϕ : A simpler and easier to implement method is estimating $\hat{\phi}$ by noting that for every τ as a function of ω (and hence the one minimizing ξ), there is a corresponding ϕ according to (3.42). Therefore, the MLE is

$$\hat{\phi} = \frac{1}{2} [(T_{2,i} + T_{3,j}) - \hat{\omega}(T_{1,i} + T_{4,j})], \quad (3.48)$$

or

$$\hat{\phi} = \frac{1}{2} [(T_{2,p} + T_{3,q}) - \hat{\omega}(T_{1,p} + T_{4,q})], \quad (3.49)$$

depending on whether $\hat{\omega}$ is given by (3.46) or (3.47). The reason for not following the same procedure as in finding $\hat{\omega}$ by using (3.43) is that the problem becomes computationally complex. First, the likelihood function assumes quite a complicated form after plugging (3.43) and (3.45) into (3.23). Second, the intersection $\hat{\phi}$ of the curves in (3.43) has to be found by solving quadratic equations with large coefficients. To be exact, $\hat{\phi}$ is the solution of

$$\begin{aligned} & 2\hat{\phi}^2[(T_{1,r} - T_{1,p}) + (T_{4,s} - T_{4,q})] + \hat{\phi}[(T_{1,p} - T_{4,q})(T_{2,r} + T_{3,s}) - \\ & (T_{1,r} - T_{4,s})(T_{2,p} + T_{3,q})] + [T_{2,p}T_{2,r}(T_{4,q} - T_{4,s}) + \\ & T_{2,p}T_{3,s}(T_{4,q} + T_{1,r}) - T_{2,r}T_{3,q}(T_{4,s} + T_{1,p}) - T_{3,q}T_{3,s}(T_{1,r} - T_{1,p})] = 0, \end{aligned}$$

where the indices p, q, r, s are the ones minimizing ξ . It has two solutions and the solution which gives

$$\hat{\phi} < \min(T_{2,i} + T_{3,j})/2, \quad i = 1, \dots, N; \quad j = 1, \dots, N,$$

is accepted to satisfy the constraints set by $I[\cdot]$ in (3.23). Hence, (3.48) or (3.49) should be chosen to estimate ϕ on the grounds of lesser computational complexity. It should be noted that $\hat{\tau}$ will be the same in both approaches when we estimate it jointly with $\hat{\omega}$ and $\hat{\phi}$ whether by expressing it in terms of ω only or in terms of ϕ only. Algorithm 4 also includes the step for estimating ϕ .

C. Summary

In this chapter, we have first derived the MLE and the CRLB for both the clock offset only and clock offset and skew cases under Gaussian noise assumption. Subsequently, the CRLB for the well-known MLE of clock offset in TPSN assuming no clock skew, under exponentially distributed delays is obtained. Afterwards, the MLEs of both the clock offset and skew for any general time synchronization protocol involving a two-way message exchange mechanism are derived assuming exponential delays and the complete algorithms used for finding these MLEs are also presented.

CHAPTER IV

COMPUTATIONALLY SIMPLIFIED SCHEMES FOR ESTIMATION OF CLOCK
OFFSET AND SKEW*

Although MLE derived in previous sections is not computationally very complex, WSNs can still benefit from some simplified schemes to estimate the clock parameters specially when the synchronization accuracy constraints are not extremely stringent but the energy conservation constraints are. In addition, for estimating both the clock offset and skew in Gaussian noise case, the knowledge of fixed portions of delay τ was required, which is usually not available beforehand. Therefore in this chapter, two simple algorithms have been proposed to estimate the clock offset and skew regardless of the distribution of the delays which are very suitable for low power demanding WSN regime. The proposed estimators can be implemented using simple steps and present remarkably low complexity. These estimators and the derived performance bounds are targeting practical applications, and are of much significance due to their robustness to the actual delay distribution involved. The main contributions of this chapter are as follows.

1. In the first scheme, first the clock skew is estimated using only the first and last data sample, since the difference between timestamps is the highest between those two for any distribution, and then maximum likelihood like estimators

*Parts of this chapter are reprinted with permissions from “Novel Clock Phase Offset and Skew Estimation Using Two-Way Timing Message Exchanges for Wireless Sensor Networks” by Kyoung-Lae Noh, Qasim M. Chaudhari, Erchin Serpedin and Bruce Suter, IEEE Transactions on Communications, Volume 55, Issue 4, April 2007 Page(s): 766 - 777 and “On Maximum Likelihood Estimation of Clock Offset and Skew in Networks With Exponential Delays”, by Qasim M. Chaudhari, Erchin Serpedin and Khalid Qaraqe, IEEE Transactions on Signal Processing, Volume 56, Issue 4, April 2008 Page(s): 1685 - 1697.

and Cramer-Rao like lower bounds are derived for the clock skew. Subsequently, the data is processed to remove the effect of skew and then the clock offset is estimated, which just requires a few computations.

2. The second scheme fits a line between two points, the differences between the first and the fourth timestamps, that are at a minimum distance apart, yielding both the clock offset and skew regardless of the underlying actual distribution.

A. Using the First and the Last Data Sample

Exploiting the fact that the clock difference between two wireless terminals is monotonically increasing (or temporary decreasing then increasing) based on the linear clock skew model adopted in this chapter, the clock difference will be maximized between the first and last timestamps. From this intuition, novel and practical clock skew estimators can be developed by using the first and last observations of timing message exchanges. Indeed, the proposed ML-Like Estimator (MLLE) maximizes the likelihood function based on the reduced set of observations (the first and last timestamps).

From (3.4), subtracting $T_{2,1}$ from $T_{2,N}$ gives

$$T_{2,N} - T_{2,1} = (T_{1,N} - T_{1,1} + X_N - X_1)\omega. \quad (4.1)$$

Similarly from (3.5), subtracting $T_{4,1}$ from $T_{4,N}$ gives

$$T_{3,N} - T_{3,1} = (T_{4,N} - T_{4,1} + Y_1 - Y_N)\omega. \quad (4.2)$$

Define the differences of the first and last timestamps as $D_{(1)} \triangleq \sum_{k=2}^N D_{1,k} = T_{1,N} - T_{1,1}$, $D_{(2)} \triangleq \sum_{k=2}^N D_{2,k} = T_{2,N} - T_{2,1}$, $D_{(3)} \triangleq \sum_{k=2}^N D_{3,k} = T_{3,N} - T_{3,1}$, and $D_{(4)} \triangleq \sum_{k=2}^N D_{4,k} = T_{4,N} - T_{4,1}$, respectively. Then (4.1) and (4.2) can be rewritten respec-

tively as

$$\begin{aligned} D_{(2)} &= (D_{(1)} + P) \omega, \\ D_{(3)} &= (D_{(4)} - R) \omega, \end{aligned}$$

where $P \triangleq X_N - X_1$ and $R \triangleq Y_N - Y_1$. Next, we analyze this simplified form to derive the MLLE and the CRLB like lower bounds for the clock skew.

1. Gaussian Delay Model

Since X_N , X_1 , Y_N , and Y_1 are i.i.d. normal distributed RVs with variance σ^2 , P and R become zero mean normal distributed RVs with variance $2\sigma^2$, respectively. Then the joint PDF of P and R is given by

$$f_{P,R}(p, r) = \frac{1}{4\pi\sigma^2} e^{-\frac{1}{4\sigma^2}(p^2+r^2)}.$$

Hence, the likelihood function becomes

$$L(\omega', \sigma^2) = \frac{1}{4\pi\sigma^2} e^{-\frac{1}{4\sigma^2}[D_{(2)}^2(\omega' - \beta)^2 + D_{(3)}^2(\omega' - \gamma)^2]},$$

where $\omega' = 1/\omega$, $\beta \triangleq D_{(1)}/D_{(2)}$ and $\gamma \triangleq D_{(4)}/D_{(3)}$. Differentiating the log-likelihood function with respect to ω' yields

$$\frac{\partial^2 \ln L(\omega', \sigma^2)}{\partial \omega'^2} = -\frac{1}{2\sigma^2} [D_{(2)}^2(\omega' - \beta) + D_{(3)}^2(\omega' - \gamma)].$$

Thus the proposed MLLE for the Gaussian delay model (GMLLE) is given by

$$\hat{\omega}_{GMLLE} = \frac{1}{\hat{\omega}'} = \frac{D_{(2)}^2 + D_{(3)}^2}{D_{(1)}D_{(2)} + D_{(3)}D_{(4)}}. \quad (4.3)$$

Again, similar procedures can be applied to derive a lower bound for the GMLLE.

The 2nd order derivative of the log likelihood function becomes

$$\frac{\partial^2 \ln L(\omega', \sigma^2)}{\partial \omega'^2} = -\frac{D_{(2)}^2 + D_{(3)}^2}{2\sigma^2}. \quad (4.4)$$

The expected value of (4.4) is given by

$$E \left[\frac{\partial^2 \ln L(\omega', \sigma^2)}{\partial \omega'^2} \right] = -\frac{E \left[D_{(2)}^2 + D_{(3)}^2 \right]}{2\sigma^2} = -\frac{D_{(1)}^2 + D_{(4)}^2 + 4\sigma^2}{2\sigma^2}.$$

Finally, the lower bound of the GMLLE is given by

$$\text{var}(\hat{\omega}_{GMLLE}) \geq \frac{\left(\frac{\partial \omega}{\partial \omega'}\right)^2}{-E \left[\frac{\partial^2 \ln L(\omega', \sigma^2)}{\partial \omega'^2} \right]} = \frac{2\sigma^2 \omega^2}{D_{(1)}^2 + D_{(4)}^2 + 4\sigma^2}. \quad (4.5)$$

Note that the complexity of the MLLEs is far less than that of the GMLE. In fact, for the GMLE, the number of required multiplications and additions are about $4N + 6$ and $10N$, respectively. While, both MLLEs require only a few multiplications and additions (less than 5) regardless of the number of beacons N . Moreover, for the GMLE, the fixed portion of delays τ must be also estimated, which requires additional computations.

2. Exponential Delay Model

For exponential delays, X_N , X_1 , Y_N , and Y_1 are assumed to be i.i.d. exponentially distributed RVs with mean α . Therefore, P and R become zero mean Laplace distributed RVs with variance $2\alpha^2$, respectively. Thus, the joint PDF of P and R is given by

$$f_{P,R}(p, r) = \left(\frac{1}{2\alpha}\right)^2 e^{-\frac{1}{\alpha}(|p|+|r|)}.$$

The likelihood function becomes

$$L(\omega, \alpha) = \left(\frac{1}{2\alpha}\right)^2 e^{-\frac{1}{\alpha} \left(\left| \frac{D_{(2)}}{\omega} - D_{(1)} \right| + \left| D_{(4)} - \frac{D_{(3)}}{\omega} \right| \right)}.$$

Substituting $1/\omega \triangleq \omega'$, the likelihood function can be rewritten as

$$L(\omega', \alpha) = \left(\frac{1}{2\alpha}\right)^2 e^{-\frac{1}{\alpha}(D_{(2)}|\omega' - \beta| + D_{(3)}|\omega' - \gamma|)},$$

where $\beta \triangleq D_{(1)}/D_{(2)}$ and $\gamma \triangleq D_{(4)}/D_{(3)}$. Then $\hat{\omega}'$ maximizing the likelihood function is given by

$$\begin{aligned} \hat{\omega}' &= \arg \min_{\omega'} (D_{(2)}|\omega' - \beta| + D_{(3)}|\omega' - \gamma|), \\ \hat{\omega}' &= \arg \min_{\omega'} \sum_{i=1}^2 K_i |\omega' - \delta_{(i)}|, \end{aligned} \quad (4.6)$$

where the order statistics $\{\delta_{(i)}\}_{i=1}^2$ are generated from the given observations $\{\beta, \gamma\}$ and K_i is the distance either ($D_{(2)}$ or $D_{(3)}$). Let $\hat{j} = \arg \min_j \sum_{i=1}^2 K_i |\delta_{(j)} - \delta_{(i)}|$, then the proposed clock skew can be derived from the solution of the equation (4.6), which is given by

$$\hat{\omega}' = \arg \min_{\omega'} \sum_{i=1}^2 K_i |\omega' - \delta_{(i)}| = \arg \min_{\omega'} h(\omega'),$$

where $h(\omega') \triangleq \sum_{i=1}^2 K_i |\omega' - \delta_{(i)}|$. Now divide the region of order statistics $\{\delta_{(i)}\}_{i=1}^2$ into 3 different regions as in Fig. 16, then the function $h(\omega')$ in the 1st region becomes

$$h(\omega') = -\sum_{i=1}^2 K_i \omega' + \sum_{i=1}^2 K_i \delta_{(i)} \quad \omega' \leq \delta_{(1)} \text{ (region 1)}.$$

Since K_i is always positive, the corresponding estimate $\hat{\omega}'$ is given by

$$\hat{\omega}' = \arg \min_{\omega'} h(\omega') = \delta_{(1)} \quad \text{(region 1)}.$$

Similarly, in the 2nd region, the function $h(\omega')$ becomes

$$h(\omega') = (K_1 - K_2) \omega' + (K_2 \delta_{(2)} - K_1 \delta_{(1)}) \quad \delta_{(1)} < \omega' \leq \delta_{(2)} \text{ (region 2)}.$$

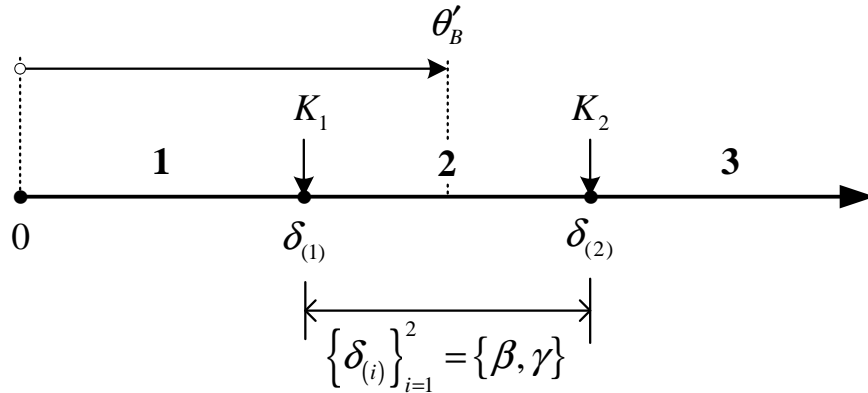


Fig. 16. Regions of the order statistics $\{\delta_{(i)}\}_{i=1}^2$.

Hence the estimate $\hat{\omega}'$ is given by

$$\hat{\omega}' = \arg \min_{\omega'} h(\omega') = \begin{cases} \delta_{(1)} & K_1 > K_2 \\ \delta_{(2)} & K_1 < K_2 \\ \text{any value} & K_1 = K_2 \end{cases} \quad \delta_{(1)} < \omega' \leq \delta_{(2)} \text{ (region 2)}.$$

Finally, the function $h(\omega')$ in the final 3rd region is given by

$$h(\omega') = \sum_{i=1}^2 K_i \omega' - \sum_{i=1}^2 K_i \delta_{(i)} \quad \delta_{(2)} < \omega' \text{ (region 3)}.$$

So the estimate $\hat{\omega}'$ in this region is

$$\hat{\omega}' = \arg \min_{\omega'} h(\omega') = \delta_{(2)} \quad \text{(region 3)}.$$

Consequently, the estimate $\hat{\omega}'$ can be determined by choosing an adequate value between the order statistics $\{\delta_{(i)}\}_{i=1}^2$. The median of $\{\delta_{(i)}\}_{i=1}^2$ maximizes the likelihood function and minimizes the mean square error of the estimate. Therefore, the MLE

of clock skew $\hat{\omega}$ for the exponential delay model is given by

$$\hat{\omega}_{EMLLE} = \begin{cases} \frac{D_2}{D_1}, & D_2 > D_3 \\ \frac{D_3}{D_4}, & D_2 < D_3 \\ \frac{1}{2} \left(\frac{D_2}{D_1} + \frac{D_4}{D_3} \right), & D_2 = D_3 \end{cases} \quad (4.7)$$

Now we are interested in the lower bound of the EMLLE to evaluate its asymptotic behavior. The derivative of the log likelihood function becomes

$$\frac{\partial \ln L(\omega', \alpha)}{\partial \omega'} = \frac{D_{(2)}}{\alpha} \text{sgn}(\omega' - \beta) + \frac{D_{(3)}}{\alpha} \text{sgn}(\omega' - \gamma). \quad (4.8)$$

Then the expected value of the square of (4.8) is given by

$$\begin{aligned} E \left[\left(\frac{\partial \ln L(\omega', \alpha)}{\partial \omega'} \right)^2 \right] &= E_{P,R} \left[\frac{D_{(2)}^2 + D_{(3)}^2 + 2D_{(2)}D_{(3)} \text{sgn}(\omega' - \beta) \text{sgn}(\omega' - \gamma)}{\alpha^2} \right] \\ &\stackrel{(c)}{=} \frac{D_{(1)}^2 + D_{(4)}^2 + 4\alpha^2}{\alpha^2}, \end{aligned}$$

where (c) is due to the fact that P and R are independent. Therefore, the lower bound of the EMLLE is given by

$$\text{var}(\hat{\omega}_{EMLLE}) \geq \frac{\left(\frac{\partial \omega}{\partial \omega'} \right)^2}{E \left[\left(\frac{\partial^2 \ln L(\omega', \alpha)}{\partial \omega'^2} \right)^2 \right]} = \frac{\alpha^2 \omega^2}{D_{(1)}^2 + D_{(4)}^2 + 4\alpha^2}. \quad (4.9)$$

In fact, we have followed the same steps used in CRLB derivation since the same reasoning and proof can be also applied to the lower bound derivation for the MLLE.

3. Combination of Clock Offset and Skew Estimation

Since the proposed MLLEs are only for estimating clock skew ω , we still need to estimate clock offset ϕ for a complete clock synchronization. Considering the given clock skew model, $T_{2,k}$ and $T_{4,k}$ are known values and ω can be estimated using the

MLLE, the sets of delay observations between two nodes can be recomposed by

$$U'_k = T_{2,k} - \hat{\omega}T_{1,k} \quad (= \tau' + \phi + X'_k), \quad (4.10)$$

$$V'_k = \hat{\omega}T_{4,k} - T_{3,k} \quad (= \tau' - \phi + Y'_k), \quad (4.11)$$

where $X'_k = \omega X_k$, $Y'_k = \omega Y_k$, and $\tau' = \omega\tau$, respectively. Notice that it can be applied to the same clock offset estimator as in (3.2) and (3.17) for Gaussian and exponential delay models, respectively. Thus, substituting the sets of delay observations gives the following clock offset estimators:

$$\hat{\phi} = \frac{\min_{1 \leq k \leq N} U'_k - \min_{1 \leq k \leq N} V'_k}{2} \quad (\text{exponential delays}), \quad (4.12)$$

$$\hat{\phi} = \frac{\overline{U'_k} - \overline{V'_k}}{2} \quad (\text{Gaussian delays}). \quad (4.13)$$

Consequently, the proposed joint clock offset and skew estimators consist of the following steps:

1. Estimate clock skew using the proposed MLLE either $\hat{\omega}_{EMLLE}$ or $\hat{\omega}_{GMLLE}$ according to the type of random delays.
2. Recompose the sets of delay observations U'_k and V'_k as shown in (4.10) and (4.11).
3. Estimate clock time offset using the estimator either (4.12) or (4.13) corresponding to the given delay model.

In fact, the proposed MLLEs require multiple message exchanges in a sync period ($N > 1$) to obtain the set of distances ($\{D_{(k)}\}_{k=1}^4$). However, these estimators can be applied not only within the same sync period, but also throughout several consecutive sync periods. In other words, a new set of observations in the next sync period can be substituted for the set of timestamps of the initial message exchange ($\{T_{k,N}\}_{k=1}^4$)

in the initial sync period. This substitution can be sequentially performed thereafter. Therefore, the proposed MLLEs can be also applied to the single message exchange model ($N = 1$) like TPSN without further modifications. The performance of the MLLEs are analyzed in the following section.

4. Simulation Results

Fig. 17 compares the variance (mean square error) of the GMLLE with the joint GMLE of clock skew and corresponding CRLB when σ is 1. It can be seen that the GMLLE performs close to the GMLE when the number of observations N is small (typically N is small in WSNs for the sake of energy efficiency), and its variance goes to zero as the number of observations increases (consistent and asymptotically efficient). Note that the GMLLE works well without the knowledge of the fixed portion of delays τ , whereas the same is required by the joint GMLE.

Fig. 18 shows the variance of the EMLLE with the joint GMLE in exponential random delay channels when α is 1. It can be seen that again the proposed MLLE is consistent and comparable to the GMLE. The consistency of the proposed MLLEs can be also checked from (4.9) and (4.5) since their lower bounds become 0 as N increases.

In order to evaluate the robustness of estimators, Fig. 19 compares the performance of the GMLE with the MLLEs in standard Gamma distributed (one of the most widely used RVs for modeling random queuing delay) random delay channels when γ is 2. Actually, both MLLEs exhibit similar performance compared to the GMLE regardless of the type of random delays. This is due to the fact that the performance of the MLLE is dominated by the set of distances ($\{D_{(k)}\}_{k=1}^4$), which does not vary much with respect to the type of random delays.

Fig. 20 compares the performance of the proposed clock offset estimator (4.13)

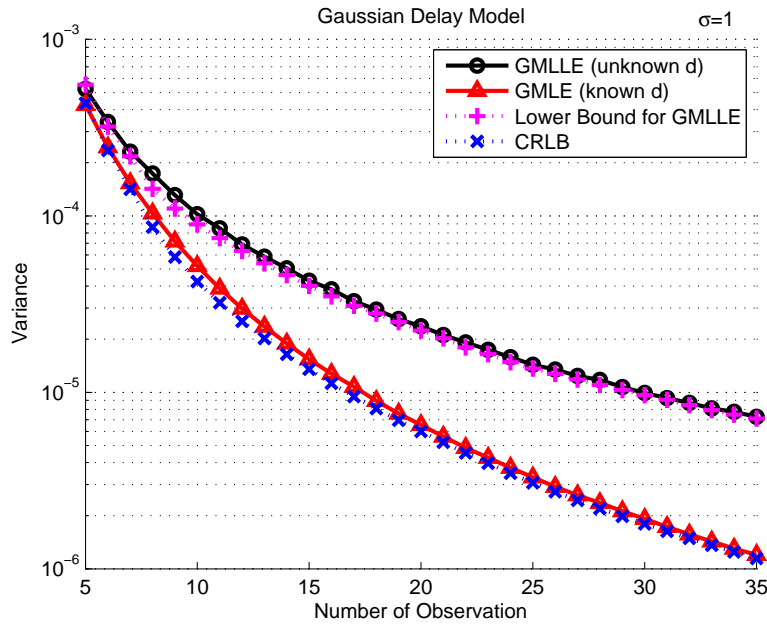


Fig. 17. Variance of the MLE of the Gaussian delay model (GMLE) and the Gaussian MLE (GMLLE) for Gaussian random delays ($\sigma = 1$).

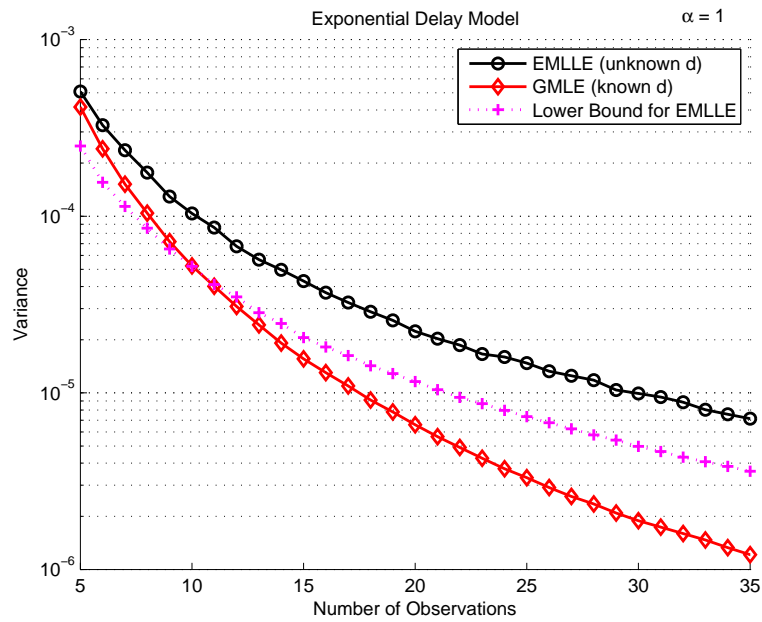


Fig. 18. Variance of the GLME and the exponential MLE (EMLLE) for exponential random delays ($\alpha = 1$).

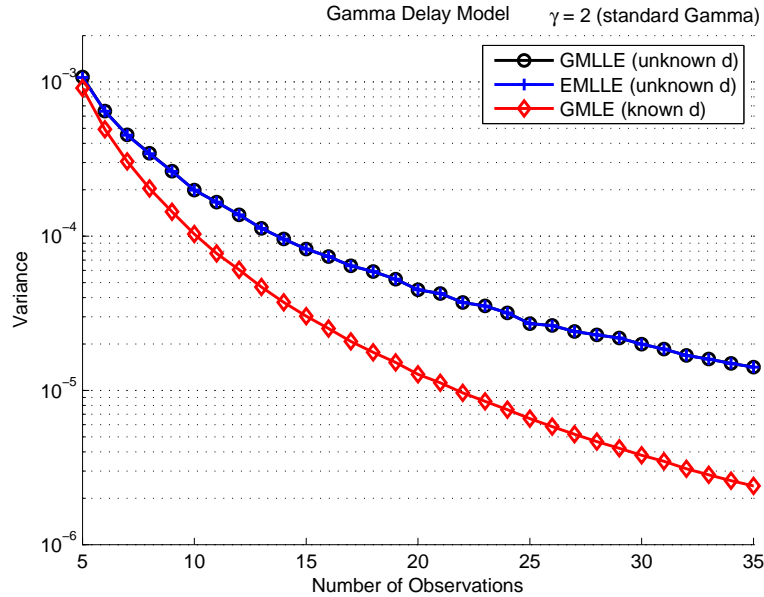


Fig. 19. Variance of the GLME and the MLEs for Gamma random delays ($\gamma = 2$).

with the joint Gaussian MLE of clock offset derived in (3.11) in Gaussian delay model when $\sigma = 0.5$. It can be seen that the joint MLE overperforms the proposed estimator due to the help of the prior knowledge of τ and the complete set of timestamps.

B. Fitting the Line Between Two Points at Minimum Distance Apart

In this section, we present an easier to implement algorithm which requires less number of computations at the expense of increased MSE, and has the most desirable feature of independence with respect to the actual delay distribution incurred. The intuition behind the idea is that (3.4) and (3.5) can be rewritten as

$$T_{2,k} = T_{1,k}\omega + \phi + (\tau + X_k)\omega,$$

$$T_{3,k} = T_{4,k}\omega + \phi - (\tau + Y_k)\omega.$$

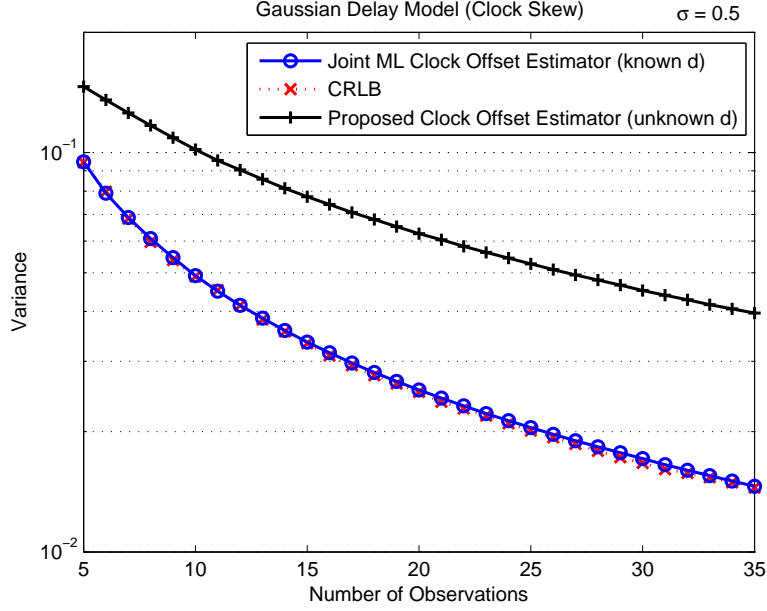


Fig. 20. Variance of the joint ML clock offset estimate and the proposed estimator for Gaussian random delays ($\sigma = 0.5$).

Notice that since ω , τ , X_k and Y_k are all positive, the points $T_{2,k}$, $k = 1, \dots, N$ will always be above the line $T_{1,k}\omega + \phi$ and the points $T_{3,k}$, $k = 1, \dots, N$ will always be below the line $T_{4,k}\omega + \phi$. Hence, a good estimate of ω and ϕ can be formed by fitting a line between the observations such that $T_{2,k}$, $k = 1, \dots, N$ are above the fitted line and $T_{3,k}$, $k = 1, \dots, N$ are below it. The strategy we have devised for a good estimate is to join the two points P_1 and P_2 , where P_1 corresponds to $\frac{1}{2} \min_{1 \leq k \leq N} \{T_{4,k} - T_{1,k}\}$ and P_2 corresponds to $\frac{1}{2} \min_{1 \leq k \leq N, k \neq i} \{T_{4,k} - T_{1,k}\}$. Representing their indices by i and j , respectively, we have

$$P_1 = \left\{ \frac{1}{2}(T_{4,i} - T_{1,i}), \frac{1}{2}(T_{2,i} + T_{3,i}) \right\},$$

and

$$P_2 = \left\{ \frac{1}{2}(T_{4,j} - T_{1,j}), \frac{1}{2}(T_{2,j} + T_{3,j}) \right\},$$

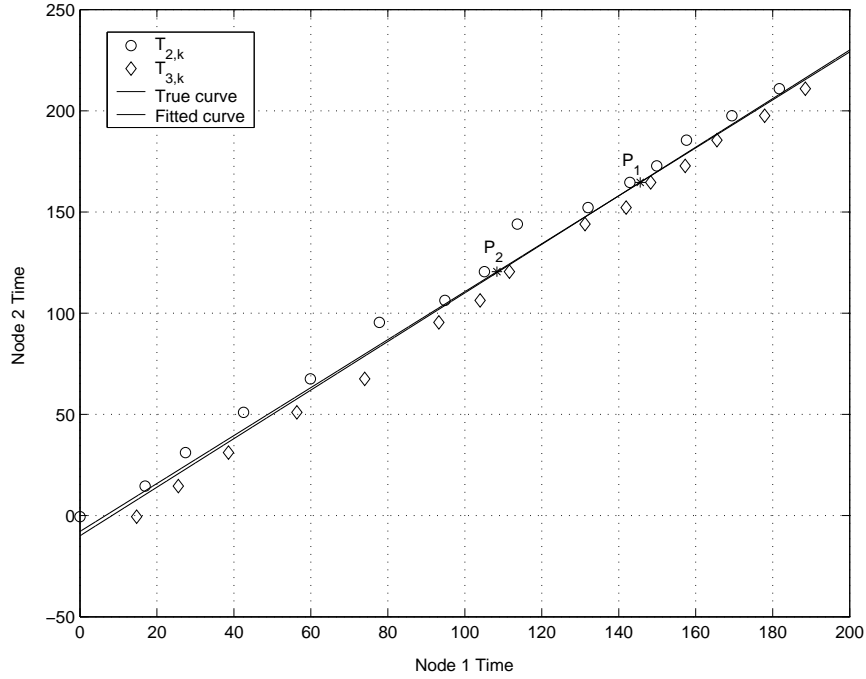


Fig. 21. The estimated fit with the original curve.

i.e., P_1 and P_2 correspond to the first two order statistics of the data set $\frac{1}{2}(T_{4,k} - T_{1,k})$, $k = 1, \dots, N$. The line formed by joining those two points is shown in Fig. 21 along with the true curve. Hence, the estimate $(\hat{\omega}, \hat{\phi})$ can be expressed as

$$\hat{\omega} = \frac{(T_{2,i} + T_{3,i})/2 - (T_{2,j} + T_{3,j})/2}{(T_{1,i} + T_{4,i})/2 - (T_{1,j} + T_{4,j})/2},$$

$$\hat{\phi} = (T_{2,i} + T_{3,i})/2 - \hat{\omega}(T_{1,i} + T_{4,i})/2.$$

When P_1 and P_2 fall very close to each other, it may happen that the fitted line exits from its boundaries and a part of it becomes either greater than some $T_{2,k}$ or less than some $T_{3,k}$. In that case, we propose to join the minimum point P_1 with one of the boundary points $\{T_{2,1}, T_{2,N}, T_{3,1}, T_{3,N}\}$ depending on which of them has the shortest distance from the initial fitted line. This algorithm is extremely simple since

Algorithm 5 Fitting the line to estimate ω and ϕ

- 1: $i = \arg \min_k \frac{1}{2} \{T_{4,k} - T_{1,k}\};$
 - 2: $j = \arg \min_{k, k \neq i} \frac{1}{2} \{T_{4,k} - T_{1,k}\};$
 - 3: $\hat{\omega} = \frac{(T_{2,i}+T_{3,i})/2 - (T_{2,j}+T_{3,j})/2}{(T_{1,i}+T_{4,i})/2 - (T_{1,j}+T_{4,j})/2}, \quad \hat{\phi} = (T_{2,i} + T_{3,i})/2 - \hat{\omega}(T_{1,i} + T_{4,i})/2;$
 - 4: **if** $(T_{2,1} < \hat{\phi} + \hat{\omega}T_{1,1})$ or $(T_{2,N} < \hat{\phi} + \hat{\omega}T_{1,N})$ **then**
 - 5: $m = \arg \min_k \{|T_{2,k} - \hat{\phi} - \hat{\omega}T_{1,k}|, |T_{3,k} - \hat{\phi} - \hat{\omega}T_{4,k}|\}, \quad k = 1, N;$
 - 6: $\hat{\omega} = \frac{(T_{2,i}+T_{3,i})/2 - (T_{2,m}+T_{3,m})/2}{(T_{1,i}+T_{4,i})/2 - (T_{1,m}+T_{4,m})/2}, \quad \hat{\phi} = (T_{2,i} + T_{3,i})/2 - \hat{\omega}(T_{1,i} + T_{4,i})/2;$
 - 7: **end if**
-

it just involves finding the first two order statistics from a set of N observations and checking the boundary conditions for the two extreme points. If the fitted line violates the boundary condition, the estimator is again formed by the same simple formula but with a point having different time index. Since this point is on the boundary, the procedure does not have to be repeated and there are no loops involved as before. The whole procedure for finding these estimates is described in Algorithm 5.

Some additional advantages of using Algorithm 5 are that ϕ can also be estimated by the y intercept of the fitted line and importantly, τ does not need to be known.

1. Simulation Results

We have simulated the performance of the MLE for fixed delay $\tau = 2$, clock offset $\phi = -10$, exponential delay parameter $\alpha = 2$ and for two different clock skews $\omega = 1.0007$ and $\omega = 1.003$. The reason of choosing different clock skews is to show a comparison of these algorithms on the performance for various actual parameters. We compare the performance of our proposed algorithm with the most general (and similar) case when (τ, ϕ, ω) have to be jointly estimated. Fig. 22 plots the Mean Square Error of both clock skew estimators for $\omega = 1.0007$ and $\omega = 1.003$ against

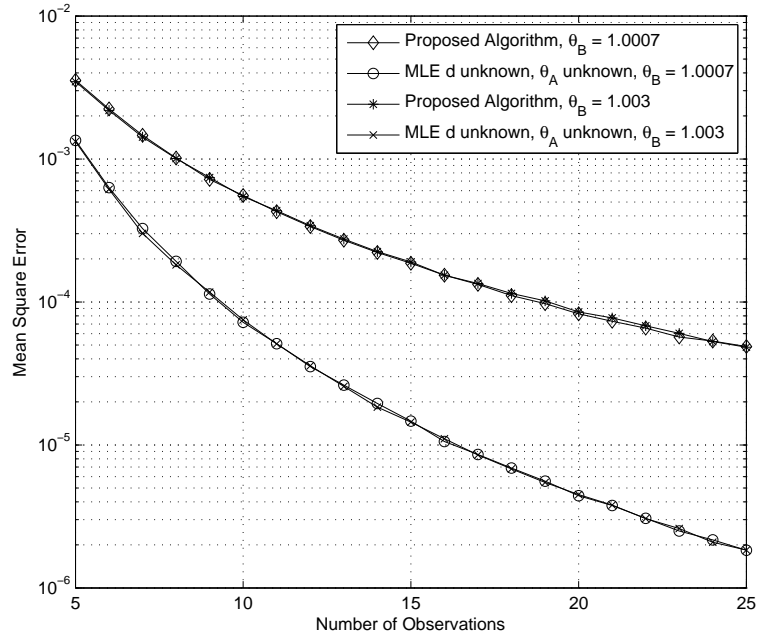


Fig. 22. Comparison of the proposed algorithm with case IV.

the number of message exchanges. It is clear from Fig. 22 that although the MLE performs better than the proposed algorithm, it can still be adopted with the sacrifice of some performance in the scenarios where energy conservation is the main issue of concern. Hence, in the light of the accuracy energy trade-off for attaining such a gain in performance by deploying MLE, we assert that the proposed algorithm is very suitable for WSNs. Moreover, there is not any significant difference between the Mean Square Error of the MLE and that of the proposed algorithm for different set of actual parameters and hence it is suited to different types of sensor nodes used today.

To check the robustness of our proposed algorithm against possible model mismatches, we have plotted the performance of the MLE in the most general Case IV and our proposed algorithm in Fig. 23 when the actual random delays come from

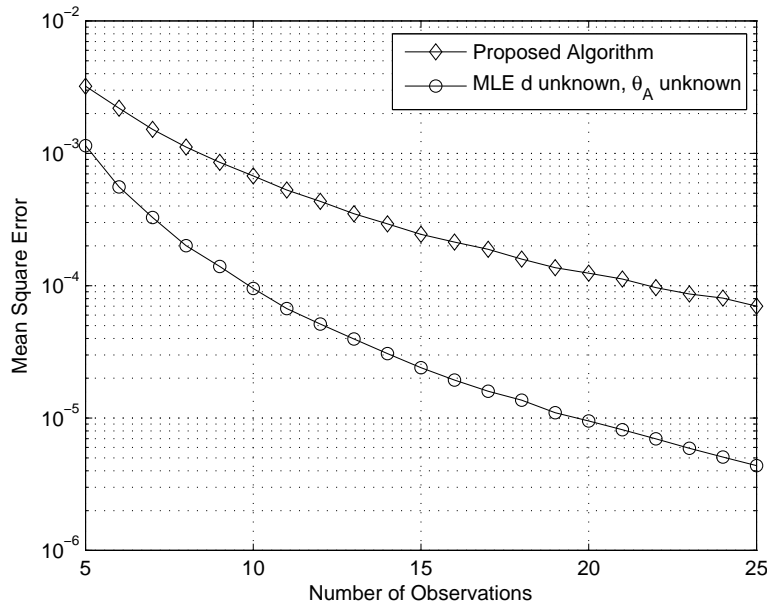


Fig. 23. Comparison of the proposed algorithm with case IV for Gamma distributed random delays.

the widely used Gamma distribution instead of the exponential distribution. Fig. 23 shows the Mean Square Error of both of these algorithms against the number of observations when the random delays were simulated as Gamma random variables with shape parameter 2 and scale parameter 1. It is interesting to observe that the difference between their performance still remains on the same scale as in Fig. 22. Therefore, the proposed algorithm is not only computationally simple and easy to implement but also robust to different environments.

2. Computational Complexity Comparison

Table III presents the number of operations required for the simplified algorithm in this section and the 4 algorithms used for deriving the MLE in Chapter III. Note that these numbers have been calculated by considering the necessary simplifications (e.g., storing the output of an operation if it is to be used later). In addition, the

operation count for Algorithm 2 and Algorithm 4 is given assuming no cycles. When their respective conditional statements become true, the code will jump around in the loop and the operation count will be multiplied by the number of cycles. Moreover, it must be kept in mind that the division is the most complex algorithm to implement in a DSP and the number of division operations must be given the highest weight while choosing between different algorithms. Finally, the operation count of our proposed algorithm is given for the worst case scenario, the probability of which is very low. For usual operation, its complexity will only be $3N + 11$ additions, $N + 4$ multiplications and 1 division.

For a comparison, observe that even for a small number of observations, e.g., 10, Algorithm 4 requires 916 additions, 205 multiplications and 200 divisions. On the other hand, the proposed algorithm requires only 61 additions, 20 multiplications and 2 divisions for 10 observations in the worst case. As the number of observations N increases, the difference between their operation counts increases significantly while the difference between their MSE decreases, making it a more viable option for large N . However, it must be remembered that in the light of the results by [31], who have reported that the energy required to transmit 1 bit over 100 meters (3 Joules) is equivalent to the energy required to execute *3 millions* of instructions, employing the MLE to achieve clock synchronization in a WSN is still a practical option.

C. Summary

In this chapter, two different computationally simple schemes have been proposed to estimate the clock parameters. The first technique utilizes the first and the last data sample to estimate the clock skew and uses this estimate for finding the clock offset. These novel ML-like estimators have been put forward for both Gaussian and

Table III. Computational complexity of each algorithm

	Additions	Multiplications	Divisions
Algorithm 1	$2N - 1$	0	N
Algorithm 2 (1 cycle)	$5N^2 + 5N$	$2N^2 + 4$	$2N^2 + 2N$
Algorithm 3	$4N^3 + 5N^2 + 2$	$2N^2$	$2N^3 + 2N^2$
Algorithm 4 (1 cycle)	$9N^2 + N + 6$	$2N^2 + 5$	$2N^2$
Proposed Algorithm	$3N + 31$	$N + 10$	2

exponential random delays and require no prior knowledge of τ . The second technique fits a line between minimum distance points and the clock offset and skew estimates are its intercept and slope, respectively. Simulation results are drawn and commented on, showing good performance by these simple estimators.

CHAPTER V

ENERGY EFFICIENT ESTIMATION OF CLOCK OFFSET FOR INACTIVE
NODES

Researchers have proposed various protocols targeting the clock synchronization in WSNs mainly based on packet synchronization techniques (see [32] - [34] for alternative schemes) which are divided into three fundamental approaches: sender-receiver synchronization (e.g., [4], [11], [35], [36], [37]), receiver-receiver synchronization (e.g., [10], [38], [39], [40]) or a hybrid of both (e.g., [13]). The two opposite requirements of closely synchronizing the network with a minimum number of RF transmissions and with high accuracy can be efficiently addressed using the approach suggested by [13], where multiple inactive nodes can hear the synchronization messages transmitted by the master node in one-way timing cells exchange mechanism. Advancing the utility of this one-way mechanism, [13] proposed the synchronization of nodes present in the communication range of the master node (broadcasting the timing beacons), where each node receiving the timing cells transmitted by the master node estimates its own clock parameters and synchronizes with the master node accordingly. However, the similar situation pertaining to the two-way timing exchange mechanism, i.e., the framework where the nodes, located in the common broadcast region of a master and slave node, can overhear the time synchronization packets between them and exploit the acquired information for achieving clock synchronization, largely remained unnoticed until [41] shed some light on it. Note that although the idea of sender-receiver synchronization is quite old and is most famously being used in NTP [1] for a long time, it is due to the wireless nature of communication channels in sensor networks that the technique of synchronization of silent nodes located in their common broadcast region can be exploited. Therefore, the clock synchronization requirements can be reason-

ably met without paying any price on the network lifetime (i.e., without exchanging additional messages for clock synchronization purposes and thereby reducing battery life) or nodes hardware (e.g., by improving the quality of the quartz crystals or by utilizing more expensive power efficient batteries). The main contributions of this chapter are as follows.

1. It derives the MLE for the clock offset of the silent nodes, which are only receiving the timing cells exchanged by the master-slave pair, and proves the uniqueness of the MLE. One very important implication of this work is that the performance of the sender-receiver protocols, whose main disadvantage has always been categorized as the high communication overhead in WSN scenarios due to their point-to-point rather than the broadcast nature, can be compared with that of receiver-receiver or hybrid protocols on equal grounds.
2. The CRLBs for the clock offsets of both the active and silent nodes are derived and used as benchmarks to assess the performance of the estimators.

A. Problem Formulation

Consider a WSN consisting of several sensor nodes as shown in Fig. 24, which dynamically elect a master node m through any master election algorithm proposed in the literature, and whose time is chosen as the *reference* time subsequently for the rest of that synchronization cycle. Depending on the sender-receiver synchronization protocol employed for operation, node m chooses another node p as the slave node at the start of the synchronization cycle. Let ϕ_p denote the clock offset of node p with respect to node m . As illustrated in Fig. 24, node m transmits timing cell 1 over the wireless channel to node p which responds by transmitting timing cell 2 to node m . The timestamps $s_j^{m \rightarrow p}$ and $r_j^{p \rightarrow m}$ are recorded by node m at pre-transmission and

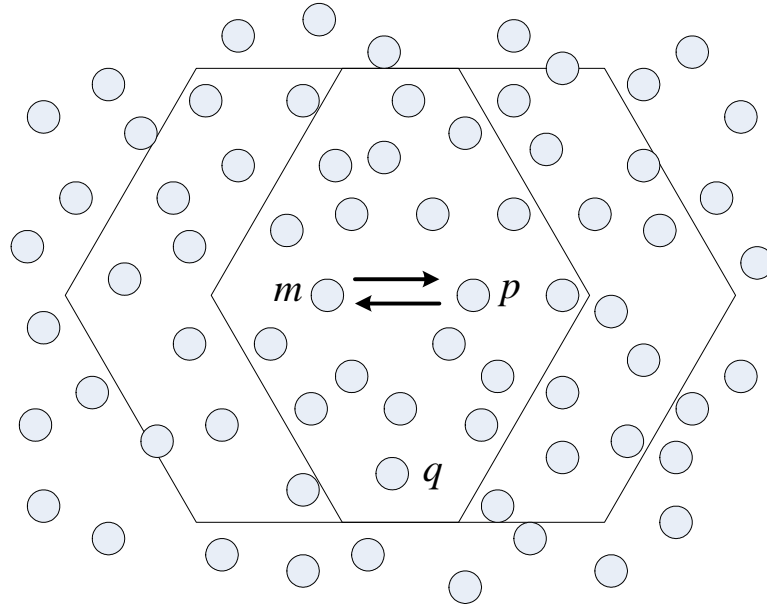


Fig. 24. A WSN with two active nodes m and p exchanging timing cells with silent nodes like q , located within the common broadcast region of the active nodes m and p .

post-reception of timing cells 1 and 2, respectively. Similarly, node p records $r_j^{m \rightarrow p}$ and $s_j^{p \rightarrow m}$ according to its own time reference (offset from node m by ϕ_p) at post-reception and pre-transmission of timing cells 1 and 2, respectively. N such timing cells are exchanged between m and p and the first of them $s_1^{m \rightarrow p}$ is chosen as the initial reference time.

Now observe from Fig. 24 that if the transmission range of sensor nodes can roughly be modeled as lying within a hexagon, then a few other nodes, e.g., node q whose clock offset with respect to node m is ϕ_q , lie within the intersection of the broadcast regions of nodes m and p . Without taking part in any communication and hence conserving considerable power, node q and other similar nodes can listen to the whole message exchange *flying through the air* between nodes m and p . For this reason, let all the transmitted messages be represented by the transmitter's index *only* without any reference to the receiving node so that $s_j^{m \rightarrow p}$ and $s_j^{p \rightarrow m}$ in the above

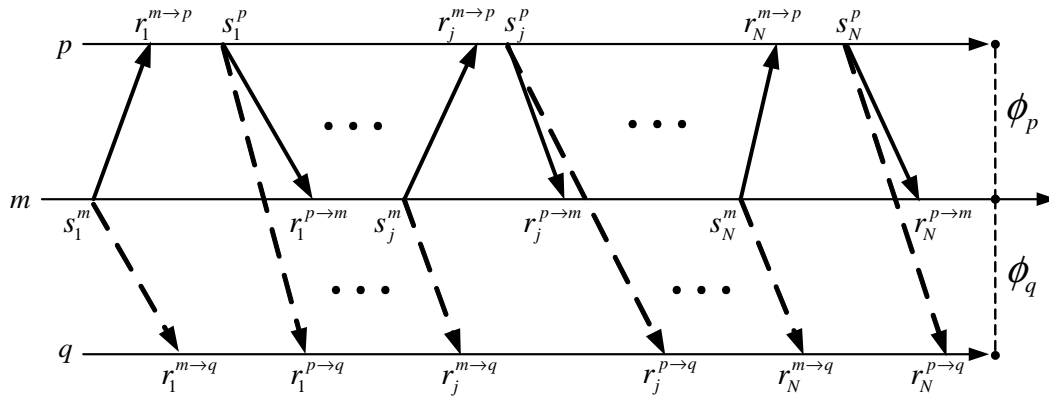


Fig. 25. A two-way timing cell exchange mechanism between nodes m and p with node q overhearing them.

paragraph now change to s_j^m and s_j^p respectively. As illustrated in Fig. 25, assume that node q timestamps the timing cells coming from nodes m and p as $r_j^{m \rightarrow q}$ and $r_j^{p \rightarrow q}$, respectively. Notice that node q is also receiving the packets $r_j^{m \rightarrow p}$, sent by node m and timestamped by node p , along with s_j^p because node p is required to send this information back to node m inside the packet containing s_j^p .

During the interval between the pre-transmission and post-reception records of a timing cell, there are different kinds of incurred link delay uncertainties in the radio message delivery, which might assume magnitudes greater than the required precision of time synchronization. Therefore, it is very important to dig deeper into the exact nature and significance of all the components comprising these sources of error. Taking into account even the minutest details, [13] classified all the link delay uncertainties incurred by the message as either deterministic or nondeterministic. The sources of delays such as send time, channel access time, interrupt handling time, receive time, etc., are nondeterministic and can range from around $5 \mu s$ to $500 ms$. On the other hand, there are deterministic sources of delays such as encoding time, transmission time, propagation time, reception time, decoding time, byte alignment

time, etc., which can range from 0 μs to 20 ms . Besides [13], numerous other authors have divided the link delay uncertainties in deterministic and nondeterministic components such as [18], [17]. Interested readers are also encouraged to go through references [15] and [16] for a detailed study of network delays and their breakdown in detail.

For the discussion in this chapter, we have assumed that the deterministic part of link delays is unknown but same for all the nodes receiving the messages from nodes m and p . This is because usually the nodes in a WSN share the same hardware specifications and characteristics and hence undergo similar transmission, reception, encoding, decoding and byte alignment times. In addition, the propagation time of RF waveforms is less than 1 μs for ranges under 300 meters which implies that for nodes lying close by at short distances from each other, the difference in the propagation time of the same message will be even less than a few nano seconds. Therefore, instead of $\tau^{m \rightarrow p}$, $\tau^{m \rightarrow q}$, or $\tau^{p \rightarrow q}$, the deterministic part of link delays is denoted as τ in this chapter.

Lastly, the nondeterministic or random link delays, $z_j^{m \rightarrow p}$, $z_j^{m \rightarrow q}$ and $z_j^{p \rightarrow q}$, have been modeled as coming from an exponential distribution with similar means. The complete discussion on the justifications behind this can be found in Chapter II.

The following equations summarize the model depicted above for $j = 1, \dots, N$.

$$\begin{aligned} r_j^{m \rightarrow p} &= s_j^m + \phi_p + \tau + z_j^{m \rightarrow p}, \\ r_j^{m \rightarrow q} &= s_j^m + \phi_q + \tau + z_j^{m \rightarrow q}, \\ r_j^{p \rightarrow q} &= s_j^p - \phi_p + \phi_q + \tau + z_j^{p \rightarrow q}, \end{aligned}$$

where $z_j^{m \rightarrow p}$, $z_j^{m \rightarrow q}$ and $z_j^{p \rightarrow q}$ are independent and identically distributed exponential random variables with the same mean λ . Rearranging the equations and introducing

the notations $U_j \triangleq r_j^{m \rightarrow p} - s_j^m$, $V_j \triangleq r_j^{m \rightarrow q} - s_j^m$ and $W_j \triangleq r_j^{p \rightarrow q} - s_j^p$ yields

$$U_j = \phi_p + \tau + z_j^{m \rightarrow p}, \quad (5.1)$$

$$V_j = \phi_q + \tau + z_j^{m \rightarrow q}, \quad (5.2)$$

$$W_j = \phi_q - \phi_p + \tau + z_j^{p \rightarrow q}. \quad (5.3)$$

Having formulated the problem and the associated model completely, next we will present a procedure for estimating the clock offsets of these silent nodes based on the ML technique at an essentially negligible cost of a few computations.

B. Maximum Likelihood Estimation

Based on the equations (5.1), (5.2) and (5.3), the likelihood function can be expressed as

$$L(\lambda, \tau, \phi_p, \phi_q) = \lambda^{-3N} \cdot \exp \left[-\frac{1}{\lambda} \sum_{j=1}^N \{U_j + V_j + W_j - 2\phi_q - 3\tau\} \right] \cdot \prod_{j=1}^N I[U_j - \phi_p - \tau] \cdot \prod_{j=1}^N I[V_j - \phi_q - \tau] \cdot \prod_{j=1}^N I[W_j - \phi_q + \phi_p - \tau], \quad (5.4)$$

where the $3N$ unit step functions $I[\cdot]$ are defined as being equal to 1 if their argument is positive and 0 otherwise, and represent the support constraints for the likelihood function. Now since these constraints do not depend on λ , the likelihood function will be maximized by $\hat{\lambda}$ for all the fixed values of (τ, ϕ_p, ϕ_q) by forcing the derivative of the log-likelihood function to be zero,

$$\frac{\partial \ln L(\lambda, \tau, \phi_p, \phi_q)}{\partial \lambda} = \frac{-3N}{\lambda} + \frac{1}{\lambda^2} \sum_{j=1}^N \{U_j + V_j + W_j - 2\phi_q - 3\tau\} = 0,$$

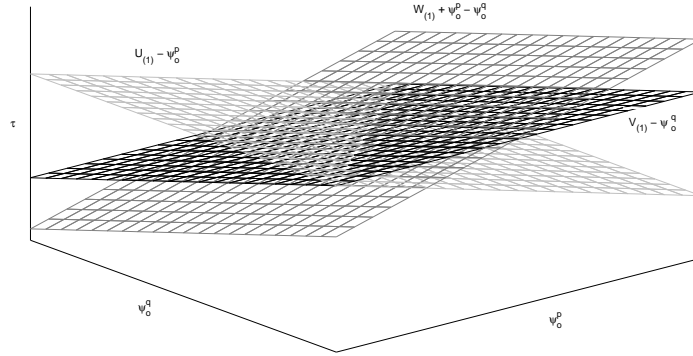


Fig. 26. Support region of the reduced likelihood function $L'(\tau, \phi_p, \phi_q)$

which implies

$$\hat{\lambda} = \frac{1}{3N} \sum_{j=1}^N \{U_j + V_j + W_j - 2\phi_q - 3\tau\}.$$

Plugging the above value of $\hat{\lambda}$ back in (5.4) and exploiting the fact that the indexed values in the unit step functions are independent of the unknown parameters yields the reduced likelihood function

$$L'(\tau, \phi_p, \phi_q) = e^{-3N} \cdot \left[\frac{1}{3N} \sum_{j=1}^N \{U_j + V_j + W_j - 2\phi_q - 3\tau\} \right]^{-3N} \cdot I[U_{(1)} - \phi_p - \tau] \cdot I[V_{(1)} - \phi_q - \tau] \cdot I[W_{(1)} - \phi_q + \phi_p - \tau], \quad (5.5)$$

where the subscript (1) denotes the minimum order statistics of the corresponding observations, i.e., $U_{(1)}$, $V_{(1)}$ and $W_{(1)}$ are the minimum values of $\{U_j\}_{j=1}^N$, $\{V_j\}_{j=1}^N$ and $\{W_j\}_{j=1}^N$, respectively.

It is clear that the reduced likelihood function $L'(\tau, \phi_p, \phi_q)$ can be maximized by minimizing the expression $\sum_{j=1}^N \{U_j + V_j + W_j - 2\phi_q - 3\tau\}$, which subsequently

becomes the cost function $f_0(\tau, \phi_q)$. Since this cost function is linear in both ϕ_q and τ , the maximum can not be found through its differentiation and hence must be searched over the boundary of its support region. Therefore, let us closely analyze this support region by writing the constraints in the following form:

$$\tau > 0, \quad (5.6)$$

$$\tau \leq U_{(1)} - \phi_p, \quad (5.7)$$

$$\tau \leq V_{(1)} - \phi_q, \quad (5.8)$$

$$\tau \leq W_{(1)} - \phi_q + \phi_p. \quad (5.9)$$

Fig. 26 draws the 3-D support region of the reduced likelihood function over which it has to be maximized, where τ is drawn as a function of ϕ_p and ϕ_q . A 2-D aerial view of this support region is drawn in Fig. 27, which illustrates the lines on the (ϕ_p, ϕ_q) plane where the intersections of the curves (5.6) - (5.9) lie. Fig. 27 is further broken down into 7 regions as shown in Fig. 28 and both of them (Figs. 27 and 28) highlight three sets of lines: solid, dashed and dotted. Each of these three sets is explained in detail in the following discussion.

- *Solid Lines:* Observe that the base of this support region is formed by the intersection of (5.6) with the surfaces $\{(5.7), (5.8), (5.9)\}$ respectively. Hence, slicing horizontally this 3-D region in Fig. 26 at $\tau = 0$ reveals the 2-D view of this base B formed by

$$B = \begin{cases} \phi_p = U_{(1)}, & -\infty < \phi_q \leq V_{(1)}, \\ \phi_q = W_{(1)} + \phi_p, & -\infty < \phi_q \leq V_{(1)}, -\infty < \phi_p \leq V_{(1)} - W_{(1)}, \\ \phi_q = V_{(1)}, & V_{(1)} - W_{(1)} \leq \phi_p \leq U_{(1)}. \end{cases} \quad (5.10)$$

The border of this base B is illustrated as solid lines in Figs. 27 and 28, and $f_0(\tau, \phi_q)$ is constrained to remain inside of it.

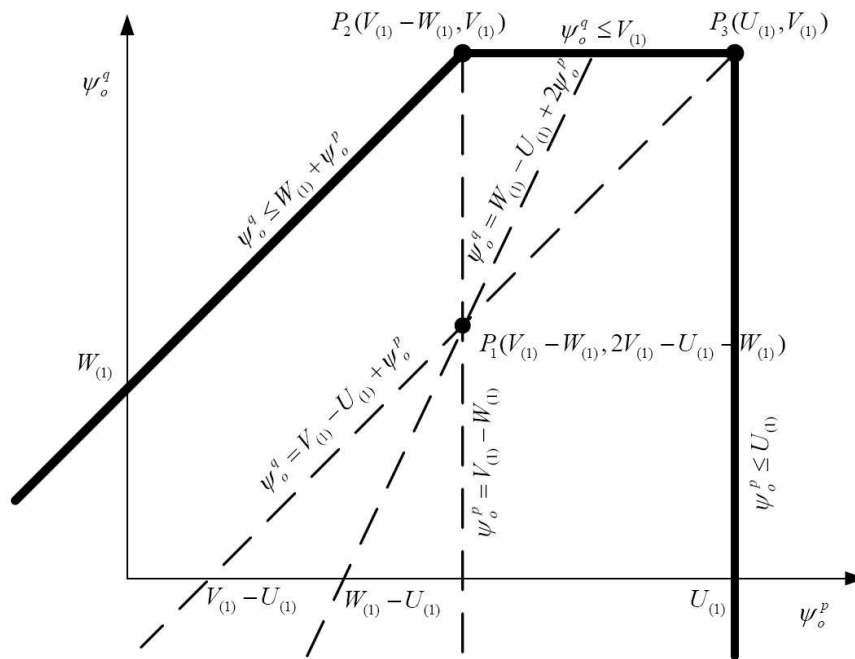


Fig. 27. An aerial 2-D view of the support region

- *Dashed Lines:* As explained above and shown in Fig. 26, the walls of the support region are formed by the three distinct surfaces (5.7), (5.8) and (5.9). The lines on the (ϕ_p, ϕ_q) plane, on which their respective intersections lie, are depicted as three dashed lines in Figs. 27 and 28 and summarized in Table IV. Also explained by this table and shown in Fig. 28, it is the point $P_1 = (V_{(1)} - W_{(1)}, 2V_{(1)} - U_{(1)} - W_{(1)})$ on (ϕ_p, ϕ_q) plane, where all the above three surfaces meet each other, that is of paramount importance for the study considered herein.
- *Dotted Lines:* For simplifying the derivation of the MLE and proving that it is unique, dotted lines are drawn in Fig. 28 in order to further break the base B into easier-to-work-with geometrical figures.

Note that in maximizing $L'(\tau, \phi_p, \phi_q)$ over the set $\Phi = \{(\tau, \phi_p, \phi_q) : \tau > 0, |\phi_p| <$

Table IV. Intersections of surfaces (5.7) - (5.9)

Surfaces	Contour of Intersection on (ϕ_p, ϕ_q) plane
(5.7) & (5.8)	Line $\phi_q = V_{(1)} - U_{(1)} + \phi_p$
(5.8) & (5.9)	Line $\phi_p = V_{(1)} - W_{(1)}$
(5.7) & (5.9)	Line $\phi_q = W_{(1)} - U_{(1)} + 2\phi_p$
(5.7) & (5.8) & (5.9)	Point $(V_{(1)} - W_{(1)}, 2V_{(1)} - U_{(1)} - W_{(1)})$

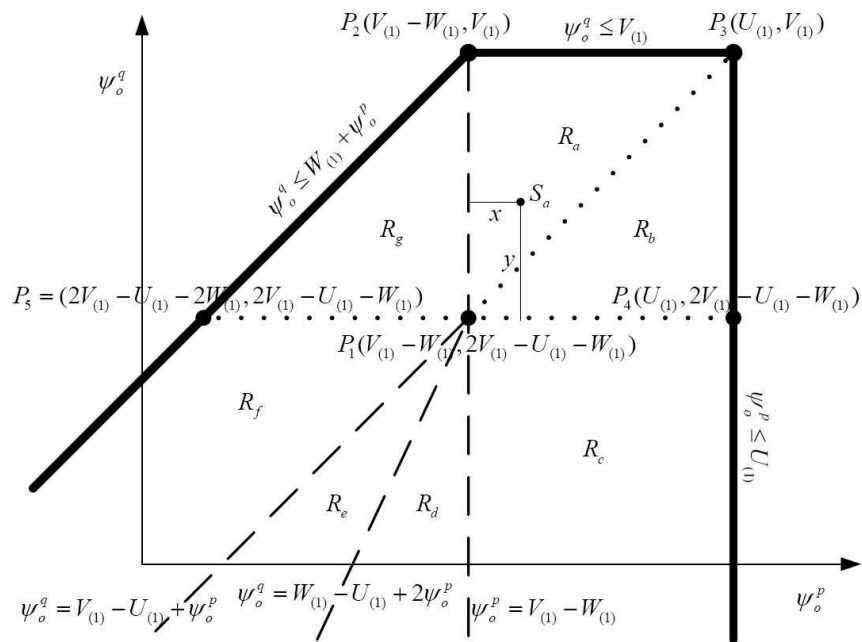


Fig. 28. A breakdown of the support region in 7 parts

$\infty, |\phi_q| < \infty\}$, four different cases need to be considered with reference to point P_1 and the point $P_3 = (U_{(1)}, V_{(1)})$ in Fig. 27.

1. $[(V_{(1)} - W_{(1)} < U_{(1)}) \cap (2V_{(1)} - U_{(1)} - W_{(1)} < V_{(1)})]$: This is the case drawn in Fig. 28 and it will suffice to derive the MLE considering it, since the other possible case is handled in a similar fashion.
2. $[(V_{(1)} - W_{(1)} > U_{(1)}) \cap (2V_{(1)} - U_{(1)} - W_{(1)} > V_{(1)})]$: In this case, boundaries of the support region and the intersections of the surfaces are drawn by a mirror image or 180° rotation of Figs. 27 and 28. The MLE remains exactly the same and its derivation follows similar arguments as in Case 1.
3. $[(V_{(1)} - W_{(1)} < U_{(1)}) \cap (2V_{(1)} - U_{(1)} - W_{(1)} > V_{(1)})]$: This case is not possible since $2V_{(1)} - U_{(1)} - W_{(1)} > V_{(1)}$ implies $V_{(1)} - W_{(1)} > U_{(1)}$, which is in contradiction with the first condition $V_{(1)} - W_{(1)} < U_{(1)}$.
4. $[(V_{(1)} - W_{(1)} > U_{(1)}) \cap (2V_{(1)} - U_{(1)} - W_{(1)} < V_{(1)})]$: This is also not possible due to a similar reason as mentioned in Case 3 above.

The following result is introduced to ease the derivation of the MLE.

Theorem 1: The MLE lies on the edge of the support region, i.e., somewhere on the ceiling of any of the surfaces (5.7) - (5.9).

Proof: Suppose that the MLE lies anywhere inside the support region at a point $C(\hat{\tau}, \hat{\phi}_p, \hat{\phi}_q)$. Now for the same $(\hat{\phi}_p, \hat{\phi}_q)$, $f_0(\hat{\tau}, \hat{\phi}_q)$ can further be minimized by increasing $\hat{\tau}$ until it touches the edge of the overlying surface.

Having considered all the possibilities for the data and having divided the base B into the regions R_a, R_b, \dots, R_g , each of these regions will be individually analyzed to derive the MLE and prove its uniqueness with the help of Theorem 1. From here onwards, to avoid labeling too many equations and hence keeping the presentation

simple, τ in inequalities (5.7), (5.8) and (5.9) will be denoted by τ_U , τ_V and τ_W , respectively.

Region R_a

Boundary Evaluation: As shown in Fig. 28, the base of the region R_a is a triangle formed by the vertices $P_1 = (V_{(1)} - W_{(1)}, 2V_{(1)} - U_{(1)} - W_{(1)})$, $P_2 = (V_{(1)} - W_{(1)}, V_{(1)})$ and $P_3 = (U_{(1)}, V_{(1)})$. To find the surface marking the boundary of this region, consider any point S_a in this region (shown in Fig. 28) whose abscissa is at distance x from abscissa(P_1) and ordinate is at distance y from ordinate(P_1). Therefore, S_a is the point with coordinates $(V_{(1)} - W_{(1)} + x, 2V_{(1)} - U_{(1)} - W_{(1)} + y)$. Notice that $x \geq 0$, $y \geq 0$ is always true since both of them are mere Euclidean distances. In addition, the relation $y \geq x$ always holds true within R_a because the point S_a lies between the lines $\phi_q = V_{(1)} - U_{(1)} + \phi_p$ and $\phi_p = V_{(1)} - W_{(1)}$. To satisfy the constraints (5.7) - (5.9) simultaneously, plug the coordinates of S_a in them such that

$$\tau_U \leq U_{(1)} - V_{(1)} + W_{(1)} - x,$$

$$\tau_V \leq U_{(1)} - V_{(1)} + W_{(1)} - y,$$

$$\tau_W \leq U_{(1)} - V_{(1)} + W_{(1)} + x - y.$$

It is clear from above that $\tau_V \leq \tau_W$ since $x \geq 0$. Also, $\tau_V \leq \tau_U$ since $y \geq x$ here. Therefore, the surface $\tau \leq V_{(1)} - \phi_q$ forms the boundary of the support region in R_a . The main points of the above discussion are summarized in Table V.

Parameter Estimation: To derive the MLE in region R_a , consider the minimization of cost function $f_0(\tau, \phi_q) = \sum_{j=1}^N \{U_j + V_j + W_j - 2\phi_q - 3\tau\}$. By virtue of Theorem 1 and the above boundary evaluation study, the MLE lies on the surface $\tau \leq V_{(1)} - \phi_q$. To see the variation in $f_0(\tau, \phi_q)$ on this surface, substitute $\tau = V_{(1)} - \phi_q$ to get the modified cost function $f'_0(\tau, \phi_q) = \sum_{j=1}^N \{U_j + V_j + W_j - 3V_{(1)} + \phi_q\}$, which depends

Table V. Boundary evaluation of region R_a

	Major Properties of Region R_a
Border	$V_{(1)} - W_{(1)} \leq \phi_p \cap V_{(1)} - U_{(1)} + \phi_p \leq \phi_q \leq V_{(1)}$
Coordinates of Point S	$(V_{(1)} - W_{(1)} + x, 2V_{(1)} - U_{(1)} - W_{(1)} + y)$
Region Characteristic	$x, y \geq 0, y \geq x$
Surfaces (5.7) - (5.9)	$\tau_U \leq U_{(1)} - V_{(1)} + W_{(1)} - x$ $\tau_V \leq U_{(1)} - V_{(1)} + W_{(1)} - y$ $\tau_W \leq U_{(1)} - V_{(1)} + W_{(1)} + x - y$
Remarks	$\tau_V \leq \tau_W$ since $x \geq 0$; $\tau_V \leq \tau_U$ since $y \geq x$
Boundary Surface	$\tau \leq V_{(1)} - \phi_q$

only on ϕ_q . It is clear that $f'_0(\tau, \phi_q)$ can be minimized by choosing $\hat{\phi}_q$ as small as possible on this particular surface, which corresponds to the point P_1 in R_a . Hence, the MLE in R_a is given by

$$\hat{\Phi}_{MLE} = \begin{bmatrix} \hat{\phi}_p \\ \hat{\phi}_q \\ \hat{\tau} \end{bmatrix} = \begin{bmatrix} V_{(1)} - W_{(1)} \\ 2V_{(1)} - U_{(1)} - W_{(1)} \\ U_{(1)} + W_{(1)} - V_{(1)} \end{bmatrix}. \quad (5.11)$$

Region R_b

Boundary Evaluation: Following the same procedure as employed for region R_a , and summarized in Table VI, it is found that $\tau \leq U_{(1)} - \phi_p$ is the boundary surface of the support region in R_b .

Parameter Estimation: According to Theorem 1, the MLE lies on the surface $\tau \leq U_{(1)} - \phi_p$. Substituting this into the cost function $f_0(\tau, \phi_q)$ yields $f'_0(\tau, \phi_q) = \sum_{j=1}^N \{U_j + V_j + W_j - 3U_{(1)} + 3\phi_p - 2\phi_q\}$. Now in this case, $f'_0(\tau, \phi_q)$ varies on the boundary surface in R_b with both ϕ_p and ϕ_q , where the minimum ϕ_p (due to the

Table VI. Boundary evaluation of region R_b

	Major Properties of Region R_b
Border	$\phi_p \leq U_{(1)} \cap 2V_{(1)} - U_{(1)} - W_{(1)} \leq \phi_q \leq V_{(1)} - U_{(1)} + \phi_p$
Coordinates of Point S	$(V_{(1)} - W_{(1)} + x, 2V_{(1)} - U_{(1)} - W_{(1)} + y)$
Region Characteristic	$x, y \geq 0, y \leq x$
Surfaces (5.7) - (5.9)	$\tau_U \leq U_{(1)} - V_{(1)} + W_{(1)} - x$ $\tau_V \leq U_{(1)} - V_{(1)} + W_{(1)} - y$ $\tau_W \leq U_{(1)} - V_{(1)} + W_{(1)} + x - y$
Remarks	$\tau_V \leq \tau_W$ since $x \geq 0$; $\tau_U \leq \tau_V$ since $y \leq x$
Boundary Surface	$\tau \leq U_{(1)} - \phi_p$

positive sign) corresponds to the point P_1 , but the maximum ϕ_q (due to the negative sign) corresponds to the point P_3 .

For deriving the MLE, consider a point S_b anywhere in the region R_b at a distance of $\sqrt{x^2 + y^2}$ from the point P_1 and with the coordinates $(V_{(1)} - W_{(1)} + x, 2V_{(1)} - U_{(1)} - W_{(1)} + y)$. It is evident that within this region, $x \geq y$. Now relating $f'_0(\tau, \phi_q)$ to the point S_b through the boundary surface yields

$$\begin{aligned}
f'_0(\tau, \phi_q) &= \sum_{j=1}^N \{U_j + V_j + W_j - 3U_{(1)} + 3(V_{(1)} - W_{(1)} + x) \\
&\quad - 2(2V_{(1)} - U_{(1)} - W_{(1)} + y)\} \\
&\propto 3x - 2y.
\end{aligned}$$

Since the maximum value y can achieve in R_b is x , it implies $3x - 2y \geq 3x - 2x = x$. Hence, the minimization problem of $f'_0(\tau, \phi_q)$ is equivalent to minimization of $3x - 2y$ which in turn is proportional to minimization of x . It is clear from Fig. 28 that x achieves its minimum value at point P_1 . It can also be verified by considering the

region R_b as a sum of vertical segments starting on the line $\phi_q = 2V_{(1)} - U_{(1)} - W_{(1)}$ and ending on the line $\phi_q = V_{(1)} - U_{(1)} + \phi_p$, having infinitesimal distances between them. Since ϕ_p is constant on each such vertical line segment, $f'_0(\tau, \phi_q)$ can be minimized by the greatest possible ϕ_q , which coincides with the line $\phi_q = V_{(1)} - U_{(1)} + \phi_p$. This gives a set of points on this line for which the minimum should be searched, which in turn can be found by noting that $f'_0(\tau, \phi_q)$ is proportional to $3\phi_p - 2\phi_p = \phi_p$ on the line $\phi_q = V_{(1)} - U_{(1)} + \phi_p$, which corresponds to the minimum ϕ_p , and hence the point P_1 .

Therefore, the MLE in R_b is the same as in R_a given by expression in (5.11).

Region R_c

Boundary Evaluation: Working along similar lines as before, Table VII summarizes the boundary evaluation problem in region R_c . The boundary surface of the support region here is enveloped by $\tau \leq U_{(1)} - \phi_p$.

Table VII. Boundary evaluation of region R_c

	Major Properties of Region R_c
Border	$V_{(1)} - W_{(1)} \leq \phi_p \leq U_{(1)} \cap \phi_q \leq 2V_{(1)} - U_{(1)} - W_{(1)}$
Coordinates of Point S	$(V_{(1)} - W_{(1)} + x, 2V_{(1)} - U_{(1)} - W_{(1)} - y)$
Region Characteristic	$x, y \geq 0,$
Surfaces (5.7) - (5.9)	$\tau_U \leq U_{(1)} - V_{(1)} + W_{(1)} - x$ $\tau_V \leq U_{(1)} - V_{(1)} + W_{(1)} + y$ $\tau_W \leq U_{(1)} - V_{(1)} + W_{(1)} + x + y$
Remarks	$\tau_V \leq \tau_W$ since $x \geq 0$; $\tau_U \leq \tau_V$ since $x, y \geq 0$
Boundary Surface	$\tau \leq U_{(1)} - \phi_p$

Parameter Estimation: Finding the MLE in region R_c is straightforward. Due to the above boundary evaluation study and Theorem 1, $f'_0(\tau, \phi_q)$ is given by $\sum_{j=1}^N \{U_j +$

$V_j + W_j - 3U_{(1)} + 3\phi_p - 2\phi_q\}$. Clearly, it can be minimized by making ϕ_p as small as possible and ϕ_q as large as possible, both of which conditions are satisfied by the point P_1 . Hence, the MLE in R_b is again given by (5.11).

Region R_d

Boundary Evaluation: As summarized in Table VIII, the boundary surface in region R_d is $\tau \leq U_{(1)} - \phi_p$.

Table VIII. Boundary evaluation of region R_d

	Major Properties of Region R_d
Border	$\phi_p \leq V_{(1)} - W_{(1)} \cap \phi_q \leq W_{(1)} - U_{(1)} + 2\phi_p$
Coordinates of Point S	$(V_{(1)} - W_{(1)} - x, 2V_{(1)} - U_{(1)} - W_{(1)} - y)$
Region Characteristic	$x, y \geq 0, y \geq 2x$
Surfaces (5.7) - (5.9)	$\tau_U \leq U_{(1)} - V_{(1)} + W_{(1)} + x$ $\tau_V \leq U_{(1)} - V_{(1)} + W_{(1)} + y$ $\tau_W \leq U_{(1)} - V_{(1)} + W_{(1)} - x + y$
Remarks	$\tau_W \leq \tau_V$ since $x \geq 0$; $\tau_U \leq \tau_W$ since $y \geq x + x \Rightarrow y - x \geq x$
Boundary Surface	$\tau \leq U_{(1)} - \phi_p$

Parameter Estimation: In region R_d again, $f'_0(\tau, \phi_q)$ is proportional to $3\phi_p - 2\phi_q$. Although the maximum ϕ_q corresponds to the point P_1 , the minimum ϕ_p does not, requiring a closer look at the region. Now consider a point S_d anywhere in R_d whose abscissa and ordinate are $V_{(1)} - W_{(1)} - x$ and $2V_{(1)} - U_{(1)} - W_{(1)} - y$, respectively. Over the point S_d and its neighborhood, $f'_0(\tau, \phi_q)$ is given by

$$\begin{aligned}
 f'_0(\tau, \phi_q) &= \sum_{j=1}^N \{U_j + V_j + W_j - 3U_{(1)} + 3(V_{(1)} - W_{(1)} - x) \\
 &\quad - 2(2V_{(1)} - U_{(1)} - W_{(1)} - y)\}, \\
 &\propto -3x + 2y.
 \end{aligned}$$

Since $y \geq 2x$ in R_d and hence the minimum value y can achieve is $2x$, $-3x + 2y \geq -3x + 2(2x) = x$. Therefore, minimization of $f'_0(\tau, \phi_q)$ corresponds to minimization of $-3x + 2y$ which subsequently requires minimization of x . Recall that $x, y \geq 0$, consequently resulting in the coordinates of point P_1 being the MLE for (ϕ_p, ϕ_q) .

Regions R_e and R_f

Boundary Evaluation: Tables IX and X show that the surface $\tau \leq W_{(1)} + \phi_p - \phi_q$ is the envelope of the support region in both R_e and R_f . Notice that these two regions could have been combined as one larger region because both the boundary surface and the MLE (as shown in the next subsection) are the same for R_f and R_e . This has not been pursued due to the difference in the boundary evaluation procedure, since $\tau_W \leq \tau_U \leq \tau_V$ in R_e , but $\tau_W \leq \tau_V \leq \tau_U$ in R_f .

Table IX. Boundary evaluation of region R_e

	Major Properties of Region R_e
Border	$W_{(1)} - U_{(1)} + 2\phi_p \leq \phi_q \leq V_{(1)} - U_{(1)} + \phi_p$
Coordinates of Point S	$(V_{(1)} - W_{(1)} - x, 2V_{(1)} - U_{(1)} - W_{(1)} - y)$
Region Characteristic	$x, y \geq 0, x \leq y \leq 2x$
Surfaces (5.7) - (5.9)	$\tau_U \leq U_{(1)} - V_{(1)} + W_{(1)} + x$ $\tau_V \leq U_{(1)} - V_{(1)} + W_{(1)} + y$ $\tau_W \leq U_{(1)} - V_{(1)} + W_{(1)} - x + y$
Remarks	$\tau_U \leq \tau_V$ since $x \leq y$; $\tau_W \leq \tau_U$ since $y \leq x + x \Rightarrow y - x \leq x$
Boundary Surface	$\tau \leq W_{(1)} + \phi_p - \phi_q$

Parameter Estimation: In these two regions, the MLE lies on the surface $\tau \leq W_{(1)} + \phi_p - \phi_q$, which is plugged into $f_0(\tau, \phi_q)$ to yield $f'_0(\tau, \phi_q) = \sum_{j=1}^N \{U_j + V_j + W_j - 3W_{(1)} - 3\phi_p + \phi_q\}$. Again, the maximum ϕ_p (owing to the negative sign) in these two regions yields the point P_1 as the solution. However, the minimum ϕ_q (owing to the

Table X. Boundary evaluation of region R_f

	Major Properties of Region R_f
Border	$V_{(1)} - U_{(1)} + \phi_p \leq \phi_q \leq W_{(1)} + \phi_p \cap \phi_q \leq 2V_{(1)} - U_{(1)} - W_{(1)}$
Coordinates of Point S	$(V_{(1)} - W_{(1)} - x, 2V_{(1)} - U_{(1)} - W_{(1)} - y)$
Region Characteristic	$x, y \geq 0, y \leq x$
Surfaces (5.7) - (5.9)	$\tau_U \leq U_{(1)} - V_{(1)} + W_{(1)} + x$ $\tau_V \leq U_{(1)} - V_{(1)} + W_{(1)} + y$ $\tau_W \leq U_{(1)} - V_{(1)} + W_{(1)} - x + y$
Remarks	$\tau_V \leq \tau_U$ since $y \leq x$; $\tau_W \leq \tau_V$ since $x \geq 0$
Boundary Surface	$\tau \leq W_{(1)} + \phi_p - \phi_q$

positive sign) corresponds to the open areas of R_e and R_f where $\phi_q \rightarrow -\infty$. Therefore, consider a point S_{ef} somewhere in any of these two regions with the coordinates $(V_{(1)} - W_{(1)} - x, 2V_{(1)} - U_{(1)} - W_{(1)} - y)$. Over this point S_{ef} and its vicinity, $f'_0(\tau, \phi_q)$ can be written as

$$\begin{aligned}
f'_0(\tau, \phi_q) &= \sum_{j=1}^N \{U_j + V_j + W_j - 3W_{(1)} - 3(V_{(1)} - W_{(1)} - x) \\
&\quad + (2V_{(1)} - U_{(1)} - W_{(1)} - y)\} \\
&\propto 3x - y.
\end{aligned}$$

Note that minimizing $f'_0(\tau, \phi_q)$ is now equivalent to minimizing the expression $3x - y$. Using the relationship $y \leq 2x$ in these two regions, x can achieve a minimum value of $y/2$ which implies $3x - y \geq 3y/2 - y = y/2$. A positive coefficient, $1/2$, with y above implies that it should be chosen as small as possible, which is achieved on point P_1 . Therefore, the MLE in these two cases is also given by the relation (5.11).

Region R_g

Boundary Evaluation: From Table XI, it is clear that the boundary surface on R_g is $\tau \leq W_{(1)} + \phi_p - \phi_q$.

Table XI. Boundary evaluation of region R_g

	Major Properties of Region R_g
Border	$\phi_p \leq V_{(1)} - W_{(1)} \cap 2V_{(1)} - U_{(1)} - W_{(1)} \leq \phi_q \leq W_{(1)} + \phi_p$
Coordinates of Point S	$(V_{(1)} - W_{(1)} - x, 2V_{(1)} - U_{(1)} - W_{(1)} + y)$
Region Characteristic	$x, y \geq 0,$
Surfaces (5.7) - (5.9)	$\tau_U \leq U_{(1)} - V_{(1)} + W_{(1)} + x$ $\tau_V \leq U_{(1)} - V_{(1)} + W_{(1)} - y$ $\tau_W \leq U_{(1)} - V_{(1)} + W_{(1)} - x - y$
Remarks	$\tau_W \leq \tau_U$ since $x, y \geq 0$; $\tau_W \leq \tau_V$ since $x \geq 0$
Boundary Surface	$\tau \leq W_{(1)} + \phi_p - \phi_q$

Parameter Estimation: In this region, the modified cost function $f'_0(\tau, \phi_q)$ is again proportional to the expression $-3\phi_p + \phi_q$. It is evident that ϕ_p should be maximized and ϕ_q should be minimized for the minimization of $f'_0(\tau, \phi_q)$, both of which can be accomplished by choosing the point P_1 .

In conclusion, the MLE $(\hat{\phi}_p, \hat{\phi}_q, \hat{\tau})$ for each region $R_a - R_g$ is given by the expressions (5.11), and hence it is unique.

In the next section, we turn our attention to deriving the Cramer-Rao Lower Bound (CRLB) for any unbiased estimator of the clock offsets ϕ_p and ϕ_q .

C. Cramer-Rao Lower Bound

In practical applications, it is extremely useful to know in advance the best performance an estimator might achieve by deriving a lower bound for it. In addition to

providing information on how well the estimator can perform, it helps the researchers in finding an unbiased estimator that has the minimum possible variance among all unbiased estimators. Also, it places a benchmark against which different estimators can be compared to rank the finest one(s), without undergoing an empirical procedure. In this particular problem, finding the CRLB is helpful for both $\hat{\phi}_p$ and $\hat{\phi}_q$. For $\hat{\phi}_q$, it can obviously set the performance benchmark for any unbiased estimator of clock offset when a node like q is silently listening to the timing cell exchange between a pair of nodes in the vicinity; whereas for $\hat{\phi}_p$, it can compare whether the clock offset of an active node like p , estimated by an inactive node like q , can perform better than the one which node p itself can estimate during a two-way timing cell exchange with reference node m using the observations s_j^m , $r_j^{m \rightarrow p}$, s_j^p and $r_j^{p \rightarrow m}$ (derived in [18]). If that is indeed the case, then any of the inactive nodes, say q , can transmit this new estimate $\hat{\phi}_p$ to node p for improved performance, albeit at the cost of one extra communication.

The CRLB theorem states that if the regularity conditions are satisfied, i.e., $E[\partial \ln L(\theta)/\partial \theta] = 0$ for all θ , the variance of any unbiased estimator $\hat{\theta}$ must satisfy the relationship

$$\text{var}(\hat{\theta}) \geq I^{-1}(\theta),$$

where $I(\theta)$ is the quantity known as Fisher Information defined as

$$I(\theta) = -E \left[\frac{\partial^2 \ln L(\theta)}{\partial \theta^2} \right] = E \left[\left(\frac{\partial \ln L(\theta)}{\partial \theta} \right)^2 \right].$$

Clearly, the domain of the likelihood function (a product of independent PDFs and hence a PDF itself) depends on both unknown parameters ϕ_p and ϕ_q due to which the order of differentiation and integration in the regularity condition can not be interchanged and hence CRLB can not be found by employing the likelihood

function. However, there is an alternative technique available for deriving the CRLB which exploits the PDF of the estimator itself as explained below.

1. CRLB for $\hat{\phi}_q$

Working on $\hat{\phi}_q$ first, note that from (5.11),

$$\begin{aligned}\hat{\phi}_q &= 2V_{(1)} - U_{(1)} - W_{(1)} \\ &= 2\left(\tau + \phi_q + z_{(1)}^{m \rightarrow q}\right) - \left(\tau + \phi_p + z_{(1)}^{m \rightarrow p}\right) - \left(\tau + \phi_q - \phi_p + z_{(1)}^{p \rightarrow q}\right) \\ &= \phi_q + 2z_{(1)}^{m \rightarrow q} - z_{(1)}^{m \rightarrow p} - z_{(1)}^{p \rightarrow q}.\end{aligned}\tag{5.12}$$

Notice that $\hat{\phi}_q$ is an unbiased function of ϕ_q , since

$$E\left[\hat{\phi}_q\right] = E\left[2V_{(1)} - U_{(1)} - W_{(1)}\right] = \phi_q + 2\frac{\lambda}{N} - \frac{\lambda}{N} - \frac{\lambda}{N} = \phi_q,$$

and its variance is

$$\text{var}\left(\hat{\phi}_q\right) = E\left[\left(\hat{\phi}_q - \phi_q\right)^2\right] = E\left[\left(2z_{(1)}^{m \rightarrow q} - z_{(1)}^{m \rightarrow p} - z_{(1)}^{p \rightarrow q}\right)^2\right] = 6\frac{\lambda^2}{N^2},\tag{5.13}$$

where the fact that the first order statistics $z_{(1)}^{m \rightarrow p}$, $z_{(1)}^{m \rightarrow q}$ and $z_{(1)}^{p \rightarrow q}$ are also exponential random variables with mean λ/N and variance λ^2/N^2 The PDF of $\hat{\phi}_q$ can be derived as follows. Consider (5.12) which can be written as

$$\hat{\phi}_q - \phi_q = 2z_{(1)}^{m \rightarrow q} - \left(z_{(1)}^{m \rightarrow p} + z_{(1)}^{p \rightarrow q}\right) = g - h,\tag{5.14}$$

where $g = 2z_{(1)}^{m \rightarrow q}$ and $h = z_{(1)}^{m \rightarrow p} + z_{(1)}^{p \rightarrow q}$ for simplicity. It is straightforward to show that the PDF of the first order statistic $z_{(1)}^{m \rightarrow p}$ from the observation set $\{z_j^{m \rightarrow p}\}_{j=1}^N$ (and

correspondingly that of $z_{(1)}^{m \rightarrow q}$ and $z_{(1)}^{p \rightarrow q}$) is given as

$$\begin{aligned} f_{z_{(1)}^{m \rightarrow p}} \left(z_{(1)}^{m \rightarrow p} \right) &= N \left[1 - F_{z_j^{m \rightarrow p}} \left(z_{(1)}^{m \rightarrow p} \right) \right]^{N-1} \cdot f_{z_j^{m \rightarrow p}} \left(z_{(1)}^{m \rightarrow p} \right) \\ &= \frac{N}{\lambda} \exp \left[-\frac{N}{\lambda} z_{(1)}^{m \rightarrow p} \right] \quad z_{(1)}^{m \rightarrow p} \geq 0, \end{aligned} \quad (5.15)$$

where $f_{z_j^{m \rightarrow p}}$ and $F_{z_j^{m \rightarrow p}}$ are the PDF and CDF of the exponential random variables $z_j^{m \rightarrow p}$, respectively. Therefore, the PDFs of the first order statistics $z_{(1)}^{m \rightarrow p}$, $z_{(1)}^{m \rightarrow q}$ and $z_{(1)}^{p \rightarrow q}$ are also exponential with mean λ/N . Now turning to (5.14) and using (5.15), it is clear that

$$f_G(g) = \frac{N}{2\lambda} \exp \left[-\frac{N}{2\lambda} g \right] I[g]. \quad (5.16)$$

Since $z_{(1)}^{m \rightarrow p}$ and $z_{(1)}^{p \rightarrow q}$ are the first order statistics of independent data sets $\{z_j^{m \rightarrow p}\}_{j=1}^N$ and $\{z_j^{p \rightarrow q}\}_{j=1}^N$, respectively, these are also independent with the distribution (5.15). To find the PDF of h , note that

$$\begin{aligned} f_H(h) &= \int_{-\infty}^{\infty} f_{z_{(1)}^{m \rightarrow p}} \left(h - z_{(1)}^{p \rightarrow q} \right) \cdot f_{z_{(1)}^{p \rightarrow q}} \left(z_{(1)}^{p \rightarrow q} \right) \cdot I \left[h - z_{(1)}^{p \rightarrow q} \right] \cdot I \left[z_{(1)}^{p \rightarrow q} \right] dz_{(1)}^{p \rightarrow q} \\ &= \frac{N^2}{\lambda^2} \int_0^h \exp \left[-\frac{N}{\lambda} \left(h - z_{(1)}^{p \rightarrow q} \right) \right] \cdot \exp \left[-\frac{N}{\lambda} z_{(1)}^{p \rightarrow q} \right] dz_{(1)}^{p \rightarrow q} \\ &= \frac{N^2}{\lambda^2} h \exp \left[-\frac{N}{\lambda} h \right] I[h], \end{aligned} \quad (5.17)$$

which is a Gamma distribution with shape parameter 2 and scale parameter λ/N . We conclude that $\hat{\phi}_q - \phi_q$ is equal to the difference between an exponential random variable and a Gamma random variable, both of which are independent and positive valued. Therefore, $g - h$ can acquire any value from $-\infty$ to ∞ and the final PDF of $\hat{\phi}_q$ can be derived using (5.16) and (5.17) as follows.

For $\hat{\phi}_q \leq \phi_q$, we have

$$\begin{aligned}
f_{\hat{\phi}_q}(\hat{\phi}_q) &= \frac{N^3}{2\lambda^3} \int_{-\infty}^{\infty} \exp\left[-\frac{N}{2\lambda}(\hat{\phi}_q - \phi_q + h)\right] I\left[\hat{\phi}_q - \phi_q + h\right] h \exp\left[-\frac{N}{\lambda}h\right] I[h] dh \\
&= \frac{N^3}{2\lambda^3} \exp\left[-\frac{N}{2\lambda}(\hat{\phi}_q - \phi_q)\right] \int_{-(\hat{\phi}_q - \phi_q)}^{\infty} h \exp\left[-\frac{3N}{2\lambda}h\right] dh \\
&= \frac{N^3}{2\lambda^3} \exp\left[-\frac{N}{2\lambda}(\hat{\phi}_q - \phi_q)\right] \left[\exp\left[\frac{3N}{2\lambda}(\hat{\phi}_q - \phi_q)\right] \left(-\frac{2\lambda(\hat{\phi}_q - \phi_q)}{3N} + \frac{4\lambda^2}{9N^2}\right) \right] \\
&= \frac{2N}{9\lambda} \exp\left[\frac{N}{\lambda}(\hat{\phi}_q - \phi_q)\right] - \frac{N^2}{3\lambda^2}(\hat{\phi}_q - \phi_q) \exp\left[\frac{N}{\lambda}(\hat{\phi}_q - \phi_q)\right] \\
&= \frac{N}{3\lambda} \left[\frac{2}{3} - \frac{N}{\lambda}(\hat{\phi}_q - \phi_q)\right] \exp\left[\frac{N}{\lambda}(\hat{\phi}_q - \phi_q)\right].
\end{aligned}$$

And for $\hat{\phi}_q \geq \phi_q$, we infer that

$$\begin{aligned}
f_{\hat{\phi}_q}(\hat{\phi}_q) &= \frac{N^3}{2\lambda^3} \int_{-\infty}^{\infty} \exp\left[-\frac{N}{2\lambda}(\hat{\phi}_q - \phi_q + h)\right] h \exp\left[-\frac{N}{\lambda}h\right] I[h] I\left[\hat{\phi}_q - \phi_q + h\right] dh \\
&= \frac{N^3}{2\lambda^3} \exp\left[-\frac{N}{2\lambda}(\hat{\phi}_q - \phi_q)\right] \int_0^{\infty} h \exp\left[-\frac{3N}{2\lambda}h\right] dh \\
&= \frac{N^3}{2\lambda^3} \exp\left[-\frac{N}{2\lambda}(\hat{\phi}_q - \phi_q)\right] \left[\frac{4\lambda^2}{9N^2}\right] \\
&= \frac{2N}{9\lambda} \exp\left[-\frac{N}{2\lambda}(\hat{\phi}_q - \phi_q)\right].
\end{aligned}$$

Therefore, the PDF of $f_{\hat{\phi}_q}(\hat{\phi}_q)$ can now be expressed as

$$f_{\hat{\phi}_q}(\hat{\phi}_q) = \begin{cases} \frac{N}{3\lambda} \left[\frac{2}{3} - \frac{N}{\lambda}(\hat{\phi}_q - \phi_q)\right] \exp\left[\frac{N}{\lambda}(\hat{\phi}_q - \phi_q)\right] & \hat{\phi}_q \leq \phi_q \\ \frac{2N}{9\lambda} \exp\left[-\frac{N}{2\lambda}(\hat{\phi}_q - \phi_q)\right] & \hat{\phi}_q \geq \phi_q \end{cases}.$$

To check if it is indeed a valid PDF, note that

$$\begin{aligned}
\frac{2N}{9\lambda} \int_{-\infty}^0 \exp\left[\frac{N}{\lambda}(\hat{\phi}_q - \phi_q)\right] d(\hat{\phi}_q - \phi_q) &= \frac{2}{9}, \\
-\frac{N^2}{3\lambda^2} \int_{-\infty}^0 (\hat{\phi}_q - \phi_q) \exp\left[\frac{N}{\lambda}(\hat{\phi}_q - \phi_q)\right] d(\hat{\phi}_q - \phi_q) &= \frac{1}{3}, \\
\frac{2N}{9\lambda} \int_0^{\infty} \exp\left[-\frac{N}{2\lambda}(\hat{\phi}_q - \phi_q)\right] d(\hat{\phi}_q - \phi_q) &= \frac{4}{9},
\end{aligned}$$

which sum up to 1. Finally, to verify its unbiasedness, note that

$$\begin{aligned}
E \left[\hat{\phi}_q \right] &= \frac{N}{3\lambda} \int_{-\infty}^{\phi_q} \hat{\phi}_q \left[\frac{2}{3} - \frac{N}{\lambda} (\hat{\phi}_q - \phi_q) \right] \exp \left[\frac{N}{\lambda} (\hat{\phi}_q - \phi_q) \right] d\hat{\phi}_q \\
&\quad + \frac{2N}{9\lambda} \int_{\phi_q}^{\infty} \hat{\phi}_q \exp \left[-\frac{N}{2\lambda} (\hat{\phi}_q - \phi_q) \right] d\hat{\phi}_q \\
&= \left(-\frac{N}{3\lambda} \phi_q^2 + \frac{2}{3} \phi_q - \frac{2\lambda}{3N} + \frac{2}{9} \phi_q + \frac{N}{3\lambda} \phi_q^2 - \frac{1}{3} \phi_q - \frac{2\lambda}{9N} \right) + \left(\frac{4}{9} \phi_q + \frac{8\lambda}{9N} \right) \\
&= \phi_q.
\end{aligned}$$

Clearly, it is not differentiable at the point $\hat{\phi}_q = \phi_q$, but exploiting its continuity at this point ($f_{\hat{\phi}_q}(\phi_q+) = f_{\hat{\phi}_q}(\phi_q-) = 2N/9\lambda$), its domain is independent of ϕ_q . Differentiating $\ln f_{\hat{\phi}_q}(\hat{\phi}_q)$ with respect to ϕ_q yields

$$\frac{\partial \ln f_{\hat{\phi}_q}(\hat{\phi}_q)}{\partial \phi_q} = \begin{cases} \frac{N}{\lambda \left[\frac{2}{3} - \frac{N}{\lambda} (\hat{\phi}_q - \phi_q) \right]} - \frac{N}{\lambda} & \hat{\phi}_q \leq \phi_q \\ \frac{N}{2\lambda} & \hat{\phi}_q \geq \phi_q \end{cases}. \quad (5.18)$$

Taking its expected value results in

$$\begin{aligned}
E \left[\frac{\partial \ln f_{\hat{\phi}_q}(\hat{\phi}_q)}{\partial \phi_q} \right] &= \int_{-\infty}^{\phi_q} \frac{N}{\lambda} \frac{N}{3\lambda} \exp \left[\frac{N}{\lambda} (\hat{\phi}_q - \phi_q) \right] d\hat{\phi}_q \\
&\quad - \int_{-\infty}^{\phi_q} \frac{N}{\lambda} \frac{N}{3\lambda} \left[\frac{2}{3} - \frac{N}{\lambda} (\hat{\phi}_q - \phi_q) \right] \exp \left[\frac{N}{\lambda} (\hat{\phi}_q - \phi_q) \right] d\hat{\phi}_q \\
&\quad + \int_{\phi_q}^{\infty} \frac{N}{2\lambda} \frac{2N}{9\lambda} \exp \left[-\frac{N}{2\lambda} (\hat{\phi}_q - \phi_q) \right] d\hat{\phi}_q \\
&= \frac{N}{3\lambda} \int_{-\infty}^{\phi_q} \left[\frac{N}{3\lambda} + \frac{N^2 (\hat{\phi}_q - \phi_q)}{\lambda^2} \right] \exp \left[\frac{N}{\lambda} (\hat{\phi}_q - \phi_q) \right] d\hat{\phi}_q \\
&\quad + \frac{N^2}{9\lambda^2} \int_{\phi_q}^{\infty} \exp \left[-\frac{N}{2\lambda} (\hat{\phi}_q - \phi_q) \right] d\hat{\phi}_q \\
&= -\frac{2N}{9\lambda} + \frac{2N}{\lambda} = 0.
\end{aligned}$$

Having satisfied both the requirements (unbiasedness and regularity condition), the license to proceed towards deriving the CRLB is available in this case now. Dif-

ferentiating (5.18) again with respect to ϕ_q ,

$$\frac{\partial^2 \ln f_{\hat{\phi}_q}(\hat{\phi}_q)}{\partial \phi_q^2} = \begin{cases} -\frac{N^2}{\lambda^2} \left[\frac{2}{3} - \frac{N}{\lambda} (\hat{\phi}_q - \phi_q) \right]^{-2} & \hat{\phi}_q \leq \phi_q \\ 0 & \hat{\phi}_q \geq \phi_q \end{cases}.$$

Taking the expectation on both sides gives

$$E \left[\frac{\partial^2 \ln f_{\hat{\phi}_q}(\hat{\phi}_q)}{\partial \phi_q^2} \right] = -\frac{N^3}{3\lambda^3} \int_{-\infty}^{\phi_q} \left[\frac{2}{3} - \frac{N}{\lambda} (\hat{\phi}_q - \phi_q) \right]^{-1} \exp \left[\frac{N}{\lambda} (\hat{\phi}_q - \phi_q) \right] d\hat{\phi}_q.$$

A change of variable $t = N/\lambda (\hat{\phi}_q - \phi_q) - 2/3$ implies

$$\begin{aligned} E \left[\frac{\partial^2 \ln f_{\hat{\phi}_q}(\hat{\phi}_q)}{\partial \phi_q^2} \right] &= \frac{N^2}{3\lambda^2} e^{2/3} \int_{-\infty}^{-2/3} t^{-1} e^t dt \\ &= \frac{N^2}{3\lambda^2} e^{2/3} \text{Ei}(-2/3) \\ &= -0.258664 \frac{N^2}{\lambda^2}, \end{aligned} \tag{5.19}$$

where $\text{Ei}(x)$ is the well known *Exponential Integral Function* defined as

$$\text{Ei}(x) = \begin{cases} -\int_{-x}^{\infty} t^{-1} e^{-t} dt = \int_{-\infty}^x t^{-1} e^t dt, & x < 0 \\ -\lim_{\epsilon \rightarrow +0} \left[\int_{-x}^{-\epsilon} t^{-1} e^{-t} dt + \int_{\epsilon}^{\infty} t^{-1} e^{-t} dt \right], & x > 0 \end{cases}.$$

In (5.19) above, the value of $\text{Ei}(-2/3)$ has been computed as -0.398409 and $e^{2/3} = 1.947734$. Therefore, CRLB for ϕ_q is given by the expression

$$\text{CRLB}(\hat{\phi}_q) = 3.866 \frac{\lambda^2}{N^2}. \tag{5.20}$$

Note that the variance of $\hat{\phi}_q$ is inversely proportional to the square of the number of observations N^2 and hence decreases very rapidly as the nodes exchange more

messages. Having derived the Fisher information, observe from (5.18) that

$$\begin{aligned} \frac{\partial \ln f_{\hat{\phi}_q}(\hat{\phi}_q)}{\partial \phi_q} &\neq I(\phi_q)(\hat{\phi}_q - \phi_q), \\ &= \frac{N^2}{3.866\lambda^2} (2V_{(1)} - U_{(1)} - W_{(1)} - \phi_q). \end{aligned}$$

Two points are worth commenting here. First, the MLE is not efficient, since it does not satisfy the relation an efficient estimator necessarily fulfills. Second, an efficient estimator for the problem targeted in this chapter does *not* exist owing to the rule: *if an efficient estimator exists, the maximum likelihood procedure will produce it*. Consequently, it is shown in Chapter VI that for symmetric delays, no unbiased estimator can be found having a lower variance than the MLE, and hence it is also the MVUE.

2. CRLB for $\hat{\phi}_p$

Since the domain of the likelihood function in (5.4) depends on ϕ_p , it can not be utilized for finding the CRLB. Working with the PDF of $\hat{\phi}_p$ using (5.11), and proceeding in a similar way as before, we have

$$\begin{aligned} \hat{\phi}_p &= V_{(1)} - W_{(1)} = \tau + \phi_q + z_{(1)}^{m \rightarrow q} - \left(\tau + \phi_q - \phi_p + z_{(1)}^{p \rightarrow q} \right), \\ &= \phi_p + z_{(1)}^{m \rightarrow q} - z_{(1)}^{p \rightarrow q}. \end{aligned}$$

The mean and variance of $\hat{\phi}_p$, are given respectively by

$$\begin{aligned} E[\hat{\phi}_p] &= E\left[\phi_p + z_{(1)}^{m \rightarrow q} - z_{(1)}^{p \rightarrow q}\right] = \phi_p + \frac{\lambda}{N} - \frac{\lambda}{N} = \phi_p, \\ E\left[\left(\hat{\phi}_p - \phi_p\right)^2\right] &= E\left[\left(z_{(1)}^{m \rightarrow q} - z_{(1)}^{p \rightarrow q}\right)^2\right] = 2\frac{\lambda^2}{N^2}. \end{aligned}$$

Having confirmed the unbiasedness of this estimator, since the difference between two exponential random variables with mean λ/N is a Laplacian random variable with

mean 0, $f_{\hat{\phi}_p}(\phi_p)$ can be written as

$$f_{\hat{\phi}_p}(\phi_p) = \begin{cases} \frac{N}{2\lambda} \exp\left[\frac{N}{\lambda}(\hat{\phi}_p - \phi_p)\right] & \hat{\phi}_p \leq \phi_p \\ \frac{N}{2\lambda} \exp\left[-\frac{N}{\lambda}(\hat{\phi}_p - \phi_p)\right] & \hat{\phi}_p \geq \phi_p \end{cases}.$$

It is evident that the PDF is symmetric around ϕ_p and hence $E[(\partial \ln f_{\hat{\phi}_p}(\phi_p)/\partial \phi_p)] = 0$. Differentiating both sides with respect to ϕ_p and taking the expectation of its square,

$$E\left[\left(\frac{\partial \ln f_{\hat{\phi}_p}(\phi_p)}{\partial \phi_p}\right)^2\right] = \frac{N^2}{\lambda^2},$$

and hence the CRLB for $\hat{\phi}_p$ can be expressed as

$$\text{CRLB}(\hat{\phi}_p) = \frac{\lambda^2}{N^2},$$

where again the variance is inversely proportional to N^2 . Since this CRLB is slightly greater than the CRLB for $\hat{\phi}_p$ derived as $\lambda^2/4N^2$ in [42], it can be concluded that instead of the silent nodes like q estimating ϕ_p and communicating this estimate to node p , node p should estimate $\hat{\phi}_p$ by itself using the two-way timing message exchange with the reference node m .

D. Simulation Results

Computer simulations have been performed to illustrate the Mean Square Error (MSE) (or variance, since the estimators are unbiased) and CRLB for the estimators $\hat{\phi}_p$ and $\hat{\phi}_q$, where the mean of the exponential link delays has been chosen as 1. Fig. 29 shows this comparison on a logarithmic scale where the MSE of both estimators decreases with the square of the number of observations. This is due to the positive only nature of the link delays justifiably modeled as exponential random

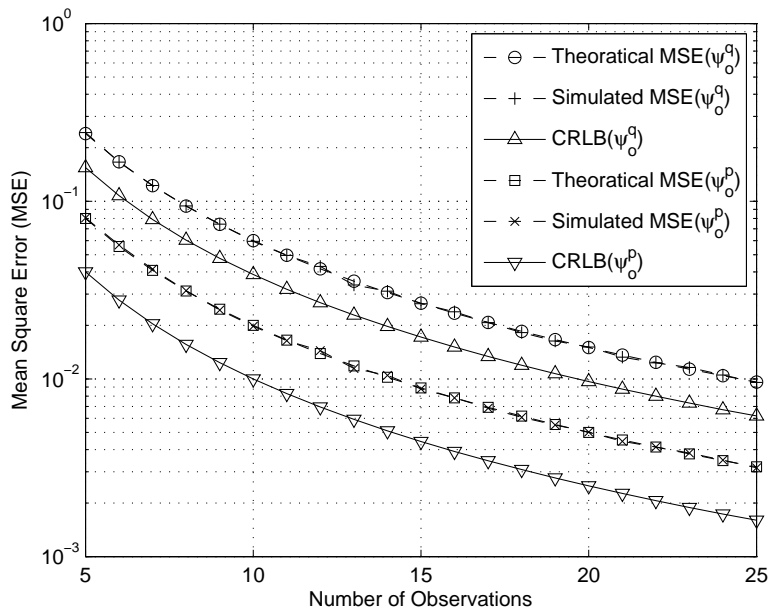


Fig. 29. Simulations illustrating the MSE of $\hat{\phi}_q$ and $\hat{\phi}_p$ on a logarithmic scale for $\lambda = 1$.

variables. Had these delays been obeying a symmetric PDF like Gaussian, the MSE would have fallen proportional to the number of data points, instead of its square. In addition, notice from Fig. 29 that there is a constant difference between the MSE of the active and inactive nodes due to the plot being drawn on a logarithmic scale. If the curves are plotted on a normal scale instead, the difference between the MSE diminishes as the number of observations increases since the clock offset estimators of both types of nodes are inversely proportional to N^2 . In conclusion, the scheme surfaces as an attractive choice more so when its cost-free quality is taken into account.

E. Summary

This chapter extends the idea based on inactive nodes overhearing a two-way timing cell exchange mechanism between the reference node and a randomly selected node by deriving the maximum likelihood estimator for the clock offset of the inactive nodes assuming the link delays obeying an exponential distribution. Since these nodes do not have to transmit any messages to other nodes, this clock synchronization approach is basically cost free with respect to energy conservation. For the clock offset estimator, Cramer-Rao lower bound has also been derived as a measure of its performance threshold. In addition, it has been proved that the clock offset estimator assumed by the active node itself, which exchanges timing cells with the reference node, performs better than the clock offset estimator assumed by a listening node. The MLE is shown not to achieve the CRLB and hence no efficient estimator exists for the concerned problem.

CHAPTER VI

SOME IMPROVED AND GENERALIZED ESTIMATION SCHEMES FOR
CLOCK SYNCHRONIZATION OF INACTIVE NODES

In the previous chapter, the MLE for the clock offset and mean link delays of the inactive node were derived under symmetric exponential delay model. This chapter not only presents better estimation techniques as compared to the MLE, but also addresses the problem under the more realistic asymmetric delay model. The main contributions of this chapter are as follows.

1. The generalized least square theory is applied on the order statistics of the synchronization packets to obtain the BLUE-OS of the clock offsets of the inactive node and the mean link delays, which is the optimal solution in the class of linear unbiased estimators.
2. The restriction of the estimates being linear is then removed to derive the MVUE of the same parameters, and no other unbiased estimator with minimum variance can be found.
3. Since the MSE can be decreased by adding a little bias to the estimator with the eventual effect of reduced variance dominating the increased bias, the Minimum Mean Square Error (MMSE) estimator with expected loss independent of clock offset and fixed delay is also obtained, thus further improving the synchronization quality.

Based on the same model with equal α , β and γ , the MLE derived in the Chapter V is expressed as

$$\hat{\Phi}_{\mathbf{S}} = \begin{bmatrix} \hat{\phi}_q \\ \hat{\phi}_p \\ \hat{\tau} \end{bmatrix} = \begin{bmatrix} 2V_{(1)} - U_{(1)} - W_{(1)} \\ V_{(1)} - W_{(1)} \\ U_{(1)} - V_{(1)} + W_{(1)} \end{bmatrix}, \quad (6.1)$$

where the subscript \mathbf{S} points to the estimates being driven for symmetric link delays and (1) denotes the minimum order statistics of their respective data sets.

A. Asymmetric Exponential Link Delays

In most communications and wireless channels, and ad-hoc networks with time-varying topologies, the network delays are asymmetric in nature. Therefore, a study for deriving the efficient estimators in this case is of paramount importance. Let the order statistics of the observations $\{U_k\}_{k=1}^N$, $\{V_k\}_{k=1}^N$ and $\{W_k\}_{k=1}^N$ be denoted as $\{U_{(k)}\}_{k=1}^N$, $\{V_{(k)}\}_{k=1}^N$ and $\{W_{(k)}\}_{k=1}^N$, respectively. Transforming the data set as

$$\begin{aligned} U'_k &\triangleq \frac{1}{\alpha} (U_k - \phi_p - \tau), \\ V'_k &\triangleq \frac{1}{\beta} (V_k - \phi_q - \tau), \\ W'_k &\triangleq \frac{1}{\gamma} (W_k - \phi_q + \phi_p - \tau), \end{aligned}$$

makes it a set of independent observations on the standardized variate and hence the distribution becomes parameter-free. The order statistics of U'_k , V'_k and W'_k are

denoted by $U'_{(k)}$, $V'_{(k)}$ and $W'_{(k)}$, respectively. Now it is straightforward to see that

$$\begin{aligned}
E [U_{(k)}] &= \phi_p + \tau + \alpha E [U'_{(k)}], \text{ var } [U_{(k)}] = \alpha^2 \text{ var } [U'_{(k)}], \\
\text{ cov } [U_{(k)} U_{(j)}] &= \alpha^2 \text{ cov } [U'_{(k)} U'_{(j)}], \\
E [V_{(k)}] &= \phi_q + \tau + \beta E [V'_{(k)}], \text{ var } [V_{(k)}] = \beta^2 \text{ var } [V'_{(k)}], \\
\text{ cov } [V_{(k)} V_{(j)}] &= \beta^2 \text{ cov } [V'_{(k)} V'_{(j)}], \\
E [W_{(k)}] &= \phi_q - \phi_p + \tau + \gamma E [W'_{(k)}], \text{ var } [W_{(k)}] = \gamma^2 \text{ var } [W'_{(k)}], \\
\text{ cov } [W_{(k)} W_{(j)}] &= \gamma^2 \text{ cov } [W'_{(k)} W'_{(j)}].
\end{aligned}$$

Next, the statistics of the ordered samples (see [26]) can be expressed as

$$\begin{aligned}
E [U'_{(k)}] &= E [V'_{(k)}] = E [W'_{(k)}] = \sum_{i=1}^k \frac{1}{(N-i+1)}, \\
\text{ var } [U'_{(k)}] &= \text{ var } [V'_{(k)}] = \text{ var } [W'_{(k)}] = \sum_{i=1}^k \frac{1}{(N-i+1)^2}, \\
\text{ cov } [U'_{(k)} U'_{(j)}] &= \text{ cov } [V'_{(k)} V'_{(j)}] = \text{ cov } [W'_{(k)} W'_{(j)}] = \sum_{i=1}^k \frac{1}{(N-i+1)^2}.
\end{aligned}$$

Therefore, the $N \times N$ symmetric positive-definite covariance matrix \mathbf{C} for each of $U'_{(k)}$, $V'_{(k)}$ and $W'_{(k)}$ takes the form:

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

The inverse of this covariance matrix can be found by the application of Gauss-

Jordan elimination as

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

Now we proceed towards estimating the clock parameters and mean link delays as follows.

1. Best Linear Unbiased Estimation Using Order Statistics

It is well known that the derivation of 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 modeling framework adopted in this chapter, the application of BLUE is not of much significance. However, for a general location-scale distribution, [25] suggested a new technique based on the derivation of BLUE using order statistics instead of just the raw observations. This technique is used in the scenario addressed in this chapter.

Let $\Phi_{\mathbf{A}} \triangleq [\phi_q \ \phi_p \ \tau \ \alpha \ \beta \ \gamma]^T$, where the subscript A denotes the relevance of estimators to asymmetric link delays and $\mathbf{z} \triangleq [U_{(1)} \ U_{(2)} \ \cdots \ U_{(N)} \ V_{(1)} \ V_{(2)} \ \cdots \ V_{(N)} \ W_{(1)} \ W_{(2)} \ \cdots \ W_{(N)}]^T$, then the linear model based on the ordered observations can

be expressed as

$$E[\mathbf{z}] = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 & \mathbf{C}_3 \end{bmatrix}^T \Phi_{\mathbf{A}} = \mathbf{Q} \Phi_{\mathbf{A}},$$

where

$$\mathbf{C}_1 = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \frac{1}{N} & \frac{1}{N} + \frac{1}{N-1} & \cdots & \sum_{k=1}^N \frac{1}{(N-k+1)} \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \quad \mathbf{C}_2 = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \frac{1}{N} & \frac{1}{N} + \frac{1}{N-1} & \cdots & \sum_{k=1}^N \frac{1}{(N-k+1)} \\ 0 & 0 & \cdots & 0 \end{bmatrix},$$

$$\mathbf{C}_3 = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ -1 & -1 & \cdots & -1 \\ 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \frac{1}{N} & \frac{1}{N} + \frac{1}{N-1} & \cdots & \sum_{k=1}^N \frac{1}{(N-k+1)} \end{bmatrix}.$$

where \mathbf{z} is a $3N \times 1$ concatenated vector of ordered data $U_{(k)}$, $V_{(k)}$ and $W_{(k)}$, \mathbf{Q} is a known matrix of dimension $3N \times 6$ and $\Phi_{\mathbf{A}}$ is the 6×1 vector of unknown parameters. Since the model has been shown to be linear in terms of the ordered observations, the BLUE can be expressed as

$$\hat{\Phi}_{\mathbf{A}} = (\mathbf{Q}^T \mathbf{C}_{\mathbf{z}}^{-1} \mathbf{Q})^{-1} \mathbf{Q}^T \mathbf{C}_{\mathbf{z}}^{-1} \mathbf{z},$$

where $\mathbf{C}_{\mathbf{z}}$ is the joint covariance matrix for $U_{(k)}$, $V_{(k)}$ and $W_{(k)}$. Due to the mutual

independence between these data sets, \mathbf{C}_z can be written as

$$\mathbf{C}_z = \begin{bmatrix} \alpha^2 \mathbf{C} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \beta^2 \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \gamma^2 \mathbf{C} \end{bmatrix},$$

and its inverse can be written as

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

Based on the above expression, it follows that

$$\mathbf{Q}^T \mathbf{C}_z^{-1} \mathbf{Q} = \begin{bmatrix} (\beta^{-2} + \gamma^{-2}) N^2 & -\gamma^{-2} N^2 & (\beta^{-2} + \gamma^{-2}) N^2 & 0 & \beta^{-2} N & \gamma^{-2} N \\ -\gamma^{-2} N & (\alpha^{-2} + \gamma^{-2}) N^2 & (\alpha^{-2} - \gamma^{-2}) N^2 & \alpha^{-2} N & 0 & -\gamma^{-2} N \\ (\beta^{-2} + \gamma^{-2}) N^2 & (\alpha^{-2} - \gamma^{-2}) N^2 & (\alpha^{-2} + \beta^{-2} + \gamma^{-2}) N^2 & \alpha^{-2} N & \beta^{-2} N & \gamma^{-2} N \\ 0 & \alpha^{-2} N & \alpha^{-2} N & \alpha^{-2} N & 0 & 0 \\ \beta^{-2} N & 0 & \beta^{-2} N & 0 & \beta^{-2} N & 0 \\ \gamma^{-2} N & -\gamma^{-2} N & \gamma^{-2} N & 0 & 0 & \gamma^{-2} N \end{bmatrix}$$

and its inverse takes the form

$$(\mathbf{Q}^T \mathbf{C}_z^{-1} \mathbf{Q})^{-1} = \begin{bmatrix} \mathbf{E}_1 & \mathbf{E}_2 \\ \mathbf{E}_3 & \mathbf{E}_4 \end{bmatrix}, \quad (6.2)$$

where

$$\mathbf{E}_1 = \frac{1}{N(N-1)} \begin{bmatrix} \alpha^2 + 4\beta^2 + \gamma^2 & 2\beta^2 + \gamma^2 & -(\alpha^2 + 2\beta^2 + \gamma^2) \\ 2\beta^2 + \gamma^2 & \beta^2 + \gamma^2 & -(\beta^2 + \gamma^2) \\ -(\alpha^2 + 2\beta^2 + \gamma^2) & -(\beta^2 + \gamma^2) & \alpha^2 + \beta^2 + \gamma^2 \end{bmatrix},$$

$$\mathbf{E}_2 = \frac{1}{N(N-1)} \begin{bmatrix} \alpha^2 & -2\beta^2 & \gamma^2 \\ 0 & -\beta^2 & \gamma^2 \\ -\alpha^2 & \beta^2 & -\gamma^2 \end{bmatrix}, \quad \mathbf{E}_3 = \frac{1}{N(N-1)} \begin{bmatrix} \alpha^2 & 0 & -\alpha^2 \\ -2\beta^2 & -\beta^2 & \beta^2 \\ \gamma^2 & \gamma^2 & -\gamma^2 \end{bmatrix},$$

$$\mathbf{E}_4 = \frac{1}{(N-1)} \begin{bmatrix} \alpha^2 & 0 & 0 \\ 0 & \beta^2 & 0 \\ 0 & 0 & \gamma^2 \end{bmatrix}.$$

Consequently,

$$(\mathbf{Q}^T \mathbf{C}_z^{-1} \mathbf{Q})^{-1} \mathbf{Q}^T \mathbf{C}_z^{-1} = \frac{1}{N(N-1)} \begin{bmatrix} \mathbf{D}_1 & \mathbf{D}_2 & \mathbf{D}_3 \end{bmatrix},$$

where the matrices \mathbf{D}_1 , \mathbf{D}_2 and \mathbf{D}_3 are defined as

$$\mathbf{D}_1 = \begin{bmatrix} -(N^2-1) & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ N^2-1 & -1 & \cdots & -1 \\ -N(N-1) & N & \cdots & N \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \quad \mathbf{D}_2 = \begin{bmatrix} 2(N^2-1) & -2 & \cdots & -2 \\ N^2-1 & -1 & \cdots & -1 \\ -(N^2-1) & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ -N(N-1) & N & \cdots & N \\ 0 & 0 & \cdots & 0 \end{bmatrix},$$

$$\mathbf{D}_3 = \begin{bmatrix} -(N^2-1) & 1 & \cdots & 1 \\ -(N^2-1) & 1 & \cdots & 1 \\ N^2-1 & -1 & \cdots & -1 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ -N(N-1) & N & \cdots & N \end{bmatrix}.$$

It follows from the above equations that

$$\begin{aligned}
\hat{\Phi}_{\mathbf{A}} &= \frac{1}{N(N-1)} \\
&\left[\begin{array}{l}
-(N^2-1)U_{(1)} + \sum_{k=2}^N U_{(k)} + 2(N^2-1)V_{(1)} - 2\sum_{k=2}^N V_{(k)} - (N^2-1)W_{(1)} + \sum_{k=2}^N W_{(k)} \\
(N^2-1)V_{(1)} - \sum_{k=2}^N V_{(k)} - (N^2-1)W_{(1)} + \sum_{k=2}^N W_{(k)} \\
(N^2-1)U_{(1)} - \sum_{k=2}^N U_{(k)} - (N^2-1)V_{(1)} + \sum_{k=2}^N V_{(k)} + (N^2-1)W_{(1)} - \sum_{k=2}^N W_{(k)} \\
-N(N-1)U_{(1)} + N\sum_{k=2}^N U_{(k)} \\
-N(N-1)V_{(1)} + N\sum_{k=2}^N V_{(k)} \\
-N(N-1)W_{(1)} + N\sum_{k=2}^N W_{(k)}
\end{array} \right] \\
&= \frac{1}{N-1} \left[\begin{array}{l}
N(2V_{(1)} - U_{(1)} - W_{(1)}) - (2\bar{V} - \bar{U} - \bar{W}) \\
N(V_{(1)} - W_{(1)}) - (\bar{V} - \bar{W}) \\
N(U_{(1)} - V_{(1)} + W_{(1)}) - (\bar{U} - \bar{V} + \bar{W}) \\
N(\bar{U} - U_{(1)}) \\
N(\bar{V} - V_{(1)}) \\
N(\bar{W} - W_{(1)})
\end{array} \right]. \quad (6.3)
\end{aligned}$$

2. Minimum Variance Unbiased Estimation

The ultimate goal in parameter estimation is often to find the estimator that achieves the minimum MSE, which explains why it is usually the criterion of performance. However, it is well known in theory that the optimal MSE estimators are not realizable in general. The MSE for an arbitrary parameter θ is given by the following expression.

$$\begin{aligned}
\text{MSE}(\hat{\theta}) &= E \left[(\hat{\theta} - \theta)^2 \right] \\
&= \text{var}(\hat{\theta}) + \text{Bias}^2(\hat{\theta}).
\end{aligned}$$

It is evident that the MSE is composed of two components, namely the estimator variance and squared bias. In the light of above, a technique chosen to attain realizable yet best estimators is to constrain the bias to be zero (since the dependence of minimum MSE 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. Hence, we proceed towards deriving the MVUE for the clock offset and mean link delays for the problem at hand.

In this chapter, the MVUE is obtained based on the Rao-Blackwell-Lehmann-Scheffé theorem, which outlines the following steps to work on. First, the likelihood function is factored according to Neymann-Fisher factorization theorem which gives the sufficient statistics \mathbf{T} . Second, the completeness of the sufficient statistics is checked. In case it is complete, any of the following two approaches yields the desired result $\hat{\theta}$ as the MVUE: either for any unbiased estimator $\check{\theta}$, $\hat{\theta} = E[\check{\theta}|\mathbf{T}]$ is evaluated, or a function $g(\mathbf{T})$ of the sufficient statistics is found such that $\hat{\theta} = g(\mathbf{T})$ is an unbiased estimator. The MVUE in the current scenario has been obtained through working on similar lines.

In the asymmetric delays case, the likelihood function for the clock offset as a function of observations $\{U_k\}_{k=1}^N$, $\{V_k\}_{k=1}^N$ and $\{W_k\}_{k=1}^N$ from (5.1), (5.2) and (5.3) is given by

$$L(\phi_q, \phi_p, \tau, \alpha, \beta, \gamma) = (\alpha\beta\gamma)^{-N} e^{-\sum_{k=1}^N \left[\frac{1}{\alpha}(U_k - \phi_p - \tau) + \frac{1}{\beta}(V_k - \phi_q - \tau) + \frac{1}{\gamma}(W_k - \phi_q + \phi_p - \tau) \right]} \times \\ I[U_{(1)} - \phi_p - \tau] I[V_{(1)} - \phi_q - \tau] I[W_{(1)} - \phi_q + \phi_p - \tau], \quad (6.4)$$

where $I[\cdot]$ denotes the unit step function. Exploiting the fact that the raw sample

mean and the ordered sample mean are actually the same, (6.4) can be factored as

$$L(\phi_q, \phi_p, \tau, \alpha, \beta, \gamma) = g_1 \left(\sum_{k=1}^N U_{(k)}, \phi_p, \alpha \right) g_2 \left(\sum_{k=1}^N V_{(k)}, \phi_q, \beta \right) g_3 \left(\sum_{k=1}^N W_{(k)}, \phi_q, \phi_p, \gamma \right) \\ g_4(U_{(1)}, \phi_p, \tau) g_5(V_{(1)}, \phi_q, \tau) g_6(W_{(1)}, \phi_q, \phi_p, \tau) h_1(\tau, \alpha, \beta, \gamma)$$

where

$$g_1 \left(\sum_{k=1}^N U_{(k)}, \phi_p, \alpha \right) = e^{-\frac{1}{\alpha} \sum_{k=1}^N (U_{(k)} - \phi_p)}, \quad g_4(U_{(1)}, \phi_p, \tau) = I[U_{(1)} - \phi_p - \tau], \\ g_2 \left(\sum_{k=1}^N V_{(k)}, \phi_q, \beta \right) = e^{-\frac{1}{\beta} \sum_{k=1}^N (V_{(k)} - \phi_q)}, \quad g_5(V_{(1)}, \phi_q, \tau) = I[V_{(1)} - \phi_q - \tau], \\ g_3 \left(\sum_{k=1}^N W_{(k)}, \phi_q, \phi_p, \gamma \right) = e^{-\frac{1}{\gamma} \sum_{k=1}^N (W_{(k)} - \phi_q + \phi_p)}, \\ g_6(W_{(1)}, \phi_q, \phi_p, \tau) = I[W_{(1)} - \phi_q + \phi_p - \tau], \quad h_1(\tau, \alpha, \beta, \gamma) = (\alpha\beta\gamma)^{-N} e^{N\tau[\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma}]}.$$

In the above relations, $g_1(\sum_{k=1}^N U_{(k)}, \phi_p, \alpha)$, $g_2(\sum_{k=1}^N V_{(k)}, \phi_q, \beta)$, $g_3(\sum_{k=1}^N W_{(k)}, \phi_q, \phi_p, \gamma)$, $g_4(U_{(1)}, \phi_p, \tau)$, $g_5(V_{(1)}, \phi_q, \tau)$ and $g_6(W_{(1)}, \phi_q, \phi_p, \tau)$ are functions depending on the data only through $\mathbf{T} = \{\sum_{k=1}^N U_{(k)}, U_{(1)}, \sum_{k=1}^N V_{(k)}, V_{(1)}, \sum_{k=1}^N W_{(k)}, W_{(1)}\}$. Therefore, according to the Neyman-Fisher factorization theorem, \mathbf{T} is a sufficient statistic for $\Phi_{\mathbf{A}}$.

Since $\dim(\mathbf{T}) = \dim(\Phi_{\mathbf{A}})$, we have to find a 6×1 vector function $\hat{\Phi}_{\mathbf{A}}$ such that $E[\hat{\Phi}_{\mathbf{A}}] = \Phi_{\mathbf{A}}$, provided that \mathbf{T} is a complete sufficient statistic. Since the pdf of \mathbf{T} is required to check if \mathbf{T} is complete, and $\sum_{k=1}^N U_{(k)}$ and $U_{(1)}$, $\sum_{k=1}^N V_{(k)}$ and $V_{(1)}$, and $\sum_{k=1}^N W_{(k)}$ and $W_{(1)}$ are not independent, we proceed as follows.

Considering into account only the data set $\{V_{(k)}\}_{k=1}^N$ first, it is evident that the pdf of the minimum order statistic $V_{(1)}$ is exponential with mean β/N , whereas the

joint pdf of $V_{(1)}, V_{(2)}, \dots, V_{(N)}$ is given by

$$f(V_{(1)}, V_{(2)}, \dots, V_{(N)}) = N! \beta^{-N} e^{-\frac{1}{\beta} \sum_{k=1}^N \{V_{(k)} - \phi_q - \tau\}} \cdot \prod_{k=1}^N I[V_{(k)} - \phi_q - \tau]. \quad (6.5)$$

Now consider the transformation as in [26],

$$\eta_k = (N - k + 1) (V_{(k)} - V_{(k-1)}), \quad k = 1, 2, \dots, N,$$

where $V_{(0)} = \phi_q + \tau$. Since $\sum_{k=1}^N (V_{(k)} - \phi_q - \tau) = \sum_{k=1}^N \eta_k$ and the Jacobian of the transformation is $N!$, a substitution in (6.5) reveals that

$$p(\eta_1, \eta_2, \dots, \eta_N) = \beta^{-N} e^{-\frac{1}{\beta} \sum_{k=1}^N \eta_k} \cdot \prod_{k=1}^N I[\eta_k],$$

i.e., η_k are *independent* exponential random variables with similar mean β . In addition, since each $\eta_k \sim \exp(\beta)$, each η_k also assumes a Gamma distribution $\eta_k \sim \Gamma(1, \beta)$. Using the relationship $\sum_{k=1}^N (V_{(k)} - V_{(1)}) = \sum_{k=2}^N \eta_k$, and the fact that each of $\eta_2, \eta_3, \dots, \eta_N$ is independent of η_1 (and hence of $V_{(1)}$, since $\eta_1 = N(V_{(1)} - \phi_q - \tau)$), $\sum_{k=1}^N (V_{(k)} - V_{(1)}) \sim \Gamma(N - 1, \beta)$ and is independent of $V_{(1)}$.

By a similar reasoning, it can be deduced that $\sum_{k=1}^N (U_{(k)} - U_{(1)}) \sim \Gamma(N - 1, \alpha)$ and $\sum_{k=1}^N (W_{(k)} - W_{(1)}) \sim \Gamma(N - 1, \gamma)$, and are independent of $U_{(1)}$ and $W_{(1)}$, respectively. Therefore, the one-to-one function $\mathbf{T}' = \{\sum_{k=1}^N (U_{(k)} - U_{(1)}), U_{(1)}, \sum_{k=1}^N (V_{(k)} - V_{(1)}), V_{(1)}, \sum_{k=1}^N (W_{(k)} - W_{(1)}), W_{(1)}\}$ of \mathbf{T} is also sufficient for estimating $\Phi_{\mathbf{A}}$ because the sufficient statistics are unique within one-to-one transformations. Consequently, \mathbf{T}' comprises of six independent random variables, which in terms of the

three-parameter Gamma distribution are given by

$$\begin{aligned}
u &= \sum_{k=1}^N (U_{(k)} - U_{(1)}) \sim \Gamma(N-1, \alpha, 0) \quad , \quad U_{(1)} \sim \Gamma(1, \alpha/N, \phi_p + \tau) \\
v &= \sum_{k=1}^N (V_{(k)} - V_{(1)}) \sim \Gamma(N-1, \beta, 0) \quad , \quad V_{(1)} \sim \Gamma(1, \beta/N, \phi_q + \tau) \\
w &= \sum_{k=1}^N (W_{(k)} - W_{(1)}) \sim \Gamma(N-1, \gamma, 0) \quad , \quad W_{(1)} \sim \Gamma(1, \gamma/N, \phi_q - \phi_p + \tau)
\end{aligned}$$

Note that the domains of u , v and w are controlled by $U_{(1)}$, $V_{(1)}$ and $W_{(1)}$, 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 $E[g(\mathbf{T}')] = \Phi_{\mathbf{A}}$. Suppose that there exists another function h for which $E[h(\mathbf{T}')] = \Phi_{\mathbf{A}}$ is also true. Then,

$$E[g(\mathbf{T}') - h(\mathbf{T}')] = E[\pi(\mathbf{T}')] = 0 \quad \forall \Phi_{\mathbf{A}}$$

where $\pi(\mathbf{T}') \triangleq g(\mathbf{T}') - h(\mathbf{T}')$ and the expectation is taken with respect to $p(\mathbf{T}'; \Phi_{\mathbf{A}})$.

As a result,

$$\begin{aligned}
&\int \int \int \int \int \int \pi(u, U_{(1)}, v, V_{(1)}, w, W_{(1)}) \cdot \frac{\alpha^{-(N-1)}}{\Gamma(N-1)} u^{N-2} e^{-\frac{u}{\alpha}} \cdot \frac{N}{\alpha} e^{-\frac{N}{\alpha}\{U_{(1)} - \phi_p - \tau\}} \\
&\frac{\beta^{-(N-1)}}{\Gamma(N-1)} v^{N-2} e^{-\frac{v}{\beta}} \cdot \frac{N}{\beta} e^{-\frac{N}{\beta}\{V_{(1)} - \phi_q - \tau\}} \cdot \frac{\gamma^{-(N-1)}}{\Gamma(N-1)} w^{N-2} e^{-\frac{w}{\gamma}} \cdot \frac{N}{\gamma} e^{-\frac{N}{\gamma}\{W_{(1)} - \phi_q + \phi_p - \tau\}} \\
&du dU_{(1)} dv dV_{(1)} dw dW_{(1)} = 0 \quad \forall \Phi_{\mathbf{A}}
\end{aligned}$$

where $R_{U_{(1)}, V_{(1)}, W_{(1)}}$ is the region defined by $I[U_{(1)} - \phi_p - \tau]$, $I[V_{(1)} - \phi_q - \tau]$ and $I[W_{(1)} - \phi_q + \phi_p - \tau]$. The above relation can be expressed as

$$\begin{aligned}
&\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \pi(u, U_{(1)}, v, V_{(1)}, w, W_{(1)}) \cdot (uvw)^{N-2} \\
&e^{-\left\{\frac{u+NU_{(1)}}{\alpha} + \frac{v+NV_{(1)}}{\beta} + \frac{w+NW_{(1)}}{\gamma}\right\}} \cdot du dU_{(1)} dv dV_{(1)} dw dW_{(1)} = 0 \quad \forall \Phi_{\mathbf{A}}
\end{aligned}$$

The expression on the left above is the six-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, leading to the conclusion that $g(\mathbf{T}') = h(\mathbf{T}')$ 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 $\Phi_{\mathbf{A}}$ when the links are asymmetric and all of α , β and γ 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 \mathfrak{R}^k , is sufficient for $\Phi_{\mathbf{A}}$ and boundedly complete, then \mathbf{T} is minimal sufficient.*

Consequently, finding an unbiased estimator for $\Phi_{\mathbf{A}}$ as a function of \mathbf{T} yields the MVUE now, according to the Rao-Blackwell-Lehmann-Scheffé theorem. Apparently, it seems difficult to find six unbiased functions of \mathbf{T} for each of ϕ_q , ϕ_p , τ , α , β and γ just by inspection. But note that ordered BLUE $\hat{\Phi}_{\mathbf{A}}$ in (6.3) is also an unbiased function of \mathbf{T} . Hence, it is concluded that the BLUE based on ordered data is also the MVUE.

$$\hat{\Phi}_{\mathbf{A}} = \begin{bmatrix} \hat{\phi}_q \\ \hat{\phi}_p \\ \hat{\tau} \\ \hat{\alpha} \\ \hat{\beta} \\ \hat{\gamma} \end{bmatrix} = \frac{1}{N-1} \begin{bmatrix} N(2V_{(1)} - U_{(1)} - W_{(1)}) - (2\bar{V} - \bar{U} - \bar{W}) \\ N(V_{(1)} - W_{(1)}) - (\bar{V} - \bar{W}) \\ N(U_{(1)} - V_{(1)} + W_{(1)}) - (\bar{U} - \bar{V} + \bar{W}) \\ N(\bar{U} - U_{(1)}) \\ N(\bar{V} - V_{(1)}) \\ N(\bar{W} - W_{(1)}) \end{bmatrix}.$$

The covariance matrix of this estimator is given by (6.2) and hence minimum variances of the clock offsets, fixed and mean delay parameters are given by its diagonal elements, whereas the total mean square error for the vector parameter $\hat{\Phi}_{\mathbf{A}}$ is the trace of this matrix.

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

$$\hat{\phi}_q = \frac{1}{(N-1)} [N(2V_{(1)} - U_{(1)} - W_{(1)}) - (2\bar{V} - \bar{U} - \bar{W})],$$

and its variance, equal to its MSE, is

$$\text{var}(\hat{\phi}_q) = \frac{1}{N(N-1)} (\alpha^2 + 4\beta^2 + \gamma^2).$$

3. Minimum Mean Square Error Estimation

Finding the MMSE estimator is not a straightforward task in any scenario, but [43] described a method to find the estimator for linear functions of the location and scale parameters with smallest mean square error among estimators with expected loss independent of the location parameters (clock offset and fixed portion of delay in the current problem). Since the derived MMSE estimator is a function of MVUE which as already been found in the previous section, we can proceed to deriving the MMSE estimator expressions for the clock offset, fixed delays and mean variable delays.

Generalizing the scalar case in [43] to a vector parameter case, for any distribution depending on location and scale parameters only, let $\mathbf{\Delta} \triangleq [\phi_q \phi_p \tau]^T$ and $\mathbf{\Xi} \triangleq [\alpha \beta \gamma]^T$. If the unique joint minimum variance unbiased estimator is denoted by $[\hat{\mathbf{\Delta}} \hat{\mathbf{\Xi}}]$, and their covariance matrix is given by (6.2), then the unique joint minimum mean square error with expected loss invariant under transformations of location and scale is

$$\begin{aligned} \hat{\mathbf{\Delta}}_{\text{MMSE}}^{\mathbf{A}} &= \hat{\mathbf{\Delta}}_{\text{MVUE}}^{\mathbf{A}} - \mathbf{E}_2 \mathbf{J}^{-1} (\mathbf{I} + \mathbf{E}_4 \mathbf{J}^{-1})^{-1} \hat{\mathbf{\Xi}}_{\text{MVUE}}^{\mathbf{A}}, \\ \hat{\mathbf{\Xi}}_{\text{MMSE}}^{\mathbf{A}} &= (\mathbf{I} + \mathbf{E}_4 \mathbf{J}^{-1})^{-1} \hat{\mathbf{\Xi}}_{\text{MVUE}}^{\mathbf{A}}, \end{aligned}$$

where

$$\mathbf{J} = \begin{bmatrix} \alpha^2 & 0 & 0 \\ 0 & \beta^2 & 0 \\ 0 & 0 & \gamma^2 \end{bmatrix}.$$

It is evident that

$$(\mathbf{I} + \mathbf{E}_4 \mathbf{J}^{-1})^{-1} = \frac{N-1}{N} \mathbf{I}$$

As a result,

$$\begin{aligned} \mathbf{E}_2 \mathbf{J}^{-1} (\mathbf{I} + \mathbf{E}_4 \mathbf{J}^{-1})^{-1} \hat{\boldsymbol{\Xi}}_{\text{MVUE}}^{\mathbf{A}} &= \\ &= \frac{1}{N(N-1)} \frac{N-1}{N} \begin{bmatrix} 1 & -2 & 1 \\ 0 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix} \frac{N}{N-1} \begin{bmatrix} \bar{U} - U_{(1)} \\ \bar{V} - V_{(1)} \\ \bar{W} - W_{(1)} \end{bmatrix} \\ &= \frac{1}{N(N-1)} \begin{bmatrix} \bar{U} - U_{(1)} - 2(\bar{V} - V_{(1)}) + \bar{W} - W_{(1)} \\ -(\bar{V} - V_{(1)}) + \bar{W} - W_{(1)} \\ -(\bar{U} - U_{(1)}) + \bar{V} - V_{(1)} - (\bar{W} - W_{(1)}) \end{bmatrix} \end{aligned}$$

Therefore, the MMSE estimators of the clock offsets, fixed delay parameter can be written as

$$\begin{aligned} \hat{\boldsymbol{\Delta}}_{\text{MMSE}}^{\mathbf{A}} &= \hat{\boldsymbol{\Delta}}_{\text{MVUE}}^{\mathbf{A}} - \mathbf{E}_2 (\mathbf{I}' + \mathbf{E}_4)^{-1} \hat{\boldsymbol{\Xi}}_{\text{MVUE}}^{\mathbf{A}}, \\ &= \frac{1}{N(N-1)} \begin{bmatrix} (N^2 - 1)(2V_{(1)} - U_{(1)} - W_{(1)}) - (N-1)(2\bar{V} - \bar{U} - \bar{W}) \\ (N^2 - 1)(V_{(1)} - W_{(1)}) - (N-1)(\bar{V} - \bar{W}) \\ (N^2 - 1)(U_{(1)} - V_{(1)} + W_{(1)}) - (N-1)(2\bar{U} - \bar{V} + \bar{W}) \end{bmatrix} \\ &= \frac{1}{N} \begin{bmatrix} (N+1)(2V_{(1)} - U_{(1)} - W_{(1)}) - (2\bar{V} - \bar{U} - \bar{W}) \\ (N+1)(V_{(1)} - W_{(1)}) - (\bar{V} - \bar{W}) \\ (N+1)(U_{(1)} - V_{(1)} + W_{(1)}) - (\bar{U} - \bar{V} + \bar{W}) \end{bmatrix}, \end{aligned}$$

where the MMSE estimator for mean variable link delays is

$$\begin{aligned} \hat{\mathbf{\Xi}}_{\text{MMSE}}^{\mathbf{A}} &= (\mathbf{I} + \mathbf{E}_4 \mathbf{J}^{-1})^{-1} \hat{\mathbf{\Xi}}_{\text{MVUE}}^{\mathbf{A}} = \frac{N-1}{N} \frac{N}{N-1} \begin{bmatrix} \bar{U} - U_{(1)} \\ \bar{V} - V_{(1)} \\ \bar{W} - W_{(1)} \end{bmatrix} \\ &= \begin{bmatrix} \bar{U} - U_{(1)} \\ \bar{V} - V_{(1)} \\ \bar{W} - W_{(1)} \end{bmatrix}. \end{aligned}$$

Therefore, the MMSE estimator for the clock offset of the inactive node is expressed as

$$\hat{\phi}_q = \frac{1}{N} [(N+1)(2V_{(1)} - U_{(1)} - W_{(1)}) - (2\bar{V} - \bar{U} - \bar{W})],$$

and its mean square error is given by

$$\text{MSE}(\hat{\phi}_q) = \frac{N+1}{N^3} (\alpha^2 + 4\beta^2 + \gamma^2),$$

which clearly outperforms the MVUE.

B. Symmetric Exponential Link Delays

The symmetric network delay assumption holds true for some realistic scenarios, e.g., when the nodes have a direct communication link between them and the topology of the network is constant. In this case, $\alpha = \beta = \gamma \triangleq \lambda$.

1. Best Linear Unbiased Estimation Using Order Statistics

Consider the BLUE based on ordered data $\mathbf{\Phi}_{\mathbf{S}} \triangleq [\phi_q \ \phi_p \ \tau \ \lambda]^T$, which is a linear function of an ordered set of observations $\{U_{(k)}\}_{k=1}^N$, $\{V_{(k)}\}_{k=1}^N$ and $\{W_{(k)}\}_{k=1}^N$. Let

$\mathbf{z} \triangleq [U_{(1)} U_{(2)} \cdots U_{(N)} V_{(1)} V_{(2)} \cdots V_{(N)} W_{(1)} W_{(2)} \cdots W_{(N)}]^T$. Then, it is evident that

$$E[\mathbf{z}] = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \mathbf{A}_3 \end{bmatrix}^T \boldsymbol{\Phi}_{\mathbf{S}} = \mathbf{Q} \boldsymbol{\Phi}_{\mathbf{S}},$$

where \mathbf{z} is the $3N \times 1$ ordered data vector, $\boldsymbol{\Phi}_{\mathbf{S}}$ is the 4×1 vector of unknown parameters and \mathbf{Q} is a known matrix of dimension $3N \times 4$ composed of

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \frac{1}{N} & \frac{1}{N} + \frac{1}{N-1} & \cdots & \sum_{k=1}^N \frac{1}{(N-k+1)} \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 \\ \frac{1}{N} & \frac{1}{N} + \frac{1}{N-1} & \cdots & \sum_{k=1}^N \frac{1}{(N-k+1)} \end{bmatrix},$$

$$\mathbf{A}_3 = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ -1 & -1 & \cdots & -1 \\ 1 & 1 & \cdots & 1 \\ \frac{1}{N} & \frac{1}{N} + \frac{1}{N-1} & \cdots & \sum_{k=1}^N \frac{1}{(N-k+1)} \end{bmatrix}.$$

The Gauss-Markov theorem yields the estimator $\hat{\boldsymbol{\Phi}}_{\mathbf{S}}$ as

$$\hat{\boldsymbol{\Phi}}_{\mathbf{S}} = (\mathbf{Q}^T \mathbf{C}_{\mathbf{z}}^{-1} \mathbf{Q})^{-1} \mathbf{Q}^T \mathbf{C}_{\mathbf{z}}^{-1} \mathbf{z}.$$

Since $\{U_{(k)}\}_{k=1}^N$, $\{V_{(k)}\}_{k=1}^N$ and $\{W_{(k)}\}_{k=1}^N$ are independent data sets, $\mathbf{C}_{\mathbf{z}}$ and its inverse $\mathbf{C}_{\mathbf{z}}^{-1}$ are now given by

$$\mathbf{C}_{\mathbf{z}} = \lambda^2 \begin{bmatrix} \mathbf{C} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{C} \end{bmatrix}, \quad \mathbf{C}_{\mathbf{z}}^{-1} = \frac{1}{\lambda^2} \begin{bmatrix} \mathbf{C}^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{C}^{-1} \end{bmatrix}.$$

It follows that

$$\mathbf{Q}^T \mathbf{C}_z^{-1} \mathbf{Q} = \frac{1}{\lambda^2} \begin{bmatrix} 0 & 2N^2 & -N^2 & 0 \\ 2N^2 & -N^2 & 2N^2 & 2N \\ 3N^2 & 0 & 2N^2 & 3N \\ 3N & 0 & 2N & 3N \end{bmatrix},$$

and its inverse is

$$(\mathbf{Q}^T \mathbf{C}_z^{-1} \mathbf{Q})^{-1} = \frac{\lambda^2}{N^2} \begin{bmatrix} 6 & 3 & -4 & 0 \\ 3 & 2 & -2 & 0 \\ -4 & -2 & \frac{9N-8}{3(N-1)} & -\frac{N}{3(N-1)} \\ 0 & 0 & -\frac{N}{3(N-1)} & \frac{N^2}{3(N-1)} \end{bmatrix}. \quad (6.6)$$

As a result,

$$(\mathbf{Q}^T \mathbf{C}_z^{-1} \mathbf{Q})^{-1} \mathbf{Q}^T \mathbf{C}_z^{-1} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 & \mathbf{B}_3 \end{bmatrix},$$

where B_1 , B_2 and B_3 are defined as

$$\mathbf{B}_1 = \begin{bmatrix} -1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 1 + \frac{1}{3N} & -\frac{1}{3N(N-1)} & \cdots & -\frac{1}{3N(N-1)} \\ -\frac{1}{3} & \frac{1}{3(N-1)} & \cdots & \frac{1}{3(N-1)} \end{bmatrix},$$

$$\mathbf{B}_2 = \begin{bmatrix} 2 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \frac{1}{3N} - 1 & -\frac{1}{3N(N-1)} & \cdots & -\frac{1}{3N(N-1)} \\ -\frac{1}{3} & \frac{1}{3(N-1)} & \cdots & \frac{1}{3(N-1)} \end{bmatrix}$$

$$\mathbf{B}_3 = \begin{bmatrix} -1 & 0 & \cdots & 0 \\ -1 & 0 & \cdots & 0 \\ 1 + \frac{1}{3N} & -\frac{1}{3N(N-1)} & \cdots & -\frac{1}{3N(N-1)} \\ -\frac{1}{3} & \frac{1}{3(N-1)} & \cdots & \frac{1}{3(N-1)} \end{bmatrix}.$$

Therefore, the BLUE using order statistics in the symmetric exponential network delays case is given by

$$\begin{aligned} \hat{\Phi}_S &= \\ &= \begin{bmatrix} 2V_{(1)} - U_{(1)} - W_{(1)} \\ V_{(1)} - W_{(1)} \\ \left(1 + \frac{1}{3N}\right) (U_{(1)} + W_{(1)}) + \left(\frac{1}{3N} - 1\right) V_{(1)} - \frac{1}{3N(N-1)} \sum_{k=2}^N (U_{(k)} + V_{(k)} + W_{(k)}) \\ -\frac{1}{3} (U_{(1)} + V_{(1)} + W_{(1)}) + \frac{1}{3(N-1)} \sum_{k=2}^N (U_{(k)} + V_{(k)} + W_{(k)}) \end{bmatrix} \\ &= \begin{bmatrix} 2V_{(1)} - U_{(1)} - W_{(1)} \\ V_{(1)} - W_{(1)} \\ \frac{1}{3(N-1)} \{3N (U_{(1)} + W_{(1)} - V_{(1)}) + 2 (2V_{(1)} - U_{(1)} - W_{(1)}) - (\bar{U} + \bar{V} + \bar{W})\} \\ \frac{N}{3(N-1)} \{(\bar{U} + \bar{V} + \bar{W}) - (U_{(1)} + V_{(1)} + W_{(1)})\} \end{bmatrix} \quad (6.7) \end{aligned}$$

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 are the same as the sample averages of ordered observations $\{U_{(k)}\}_{k=1}^N$ and $\{V_{(k)}\}_{k=1}^N$. Note that the BLUE of the clock offset based on order statistics matches the MLE of the clock offset in (6.1).

2. Minimum Variance Unbiased Estimation

In the symmetric case when $\alpha = \beta = \gamma \triangleq \lambda$, the likelihood function for the clock offset as a function of observations $\{U_k\}_{k=1}^N$, $\{V_k\}_{k=1}^N$ and $\{W_k\}_{k=1}^N$ from (5.1), (5.2)

and (5.3) can be expressed as

$$L(\phi_q, \phi_p, \tau, \lambda) = \lambda^{-3N} \cdot e^{-\frac{1}{\lambda} \sum_{k=1}^N [U_k + V_k + W_k - 2\phi_q - 3\tau]} \cdot I[U_{(1)} - \phi_p - \tau] \cdot I[V_{(1)} - \phi_q - \tau] \cdot I[W_{(1)} - \phi_q + \phi_p - \tau].$$

Note that due to the equality of the sample mean for both raw and ordered observations, the likelihood function above can be factored as

$$L(\phi_q, \phi_p, \tau, \lambda) = g_1 \left(\sum_{k=1}^N U_{(k)}, \sum_{k=1}^N V_{(k)}, \sum_{k=1}^N W_{(k)}, \lambda \right) g_2(U_{(1)}, \phi_p, \tau) \cdot g_3(V_{(1)}, \phi_q, \tau) g_4(W_{(1)}, \phi_q, \phi_p, \tau) h_1(\phi_q, \tau, \lambda),$$

where

$$g_1 \left(\sum_{k=1}^N U_{(k)}, \sum_{k=1}^N V_{(k)}, \sum_{k=1}^N W_{(k)}, \lambda \right) = e^{-\frac{1}{\lambda} \sum_{k=1}^N [U_{(k)} + V_{(k)} + W_{(k)}]},$$

$$g_2(U_{(1)}, \phi_p, \tau) = I[U_{(1)} - \phi_p - \tau], g_3(V_{(1)}, \phi_q, \tau) = I[V_{(1)} - \phi_q - \tau],$$

$$g_4(W_{(1)}, \phi_q, \phi_p, \tau) = I[W_{(1)} - \phi_q + \phi_p - \tau], h_1(\phi_q, \tau, \lambda) = \lambda^{-3N} e^{\frac{N}{\lambda} [2\phi_q + 3\tau]}.$$

It is evident that $\mathbf{T} = \{\sum_{k=1}^N (U_{(k)} + V_{(k)} + W_{(k)}), U_{(1)}, V_{(1)}, W_{(1)}\}$ are the minimal sufficient statistic according to Neymann-Fisher Factorization theorem. Now proceeding similarly as before, $\sum_{k=1}^N (U_{(k)} + V_{(k)} + W_{(k)})$ is dependent on $U_{(1)}, V_{(1)}$ and $W_{(1)}$. As a result, \mathbf{T} can be transformed into $\mathbf{T}' = \{\sum_{k=1}^N (U_{(k)} - U_{(1)} + V_{(k)} - V_{(1)} + W_{(k)} - W_{(1)}), U_{(1)}, V_{(1)}, W_{(1)}\}$. It can be concluded from the discussion in the last section that $\sum_{k=1}^N (U_{(k)} - U_{(1)} + V_{(k)} - V_{(1)} + W_{(k)} - W_{(1)})$ is Gamma distributed with parameters $(3(N-1), \lambda)$. Hence, \mathbf{T}' is a combination of four independent random variables, which in terms of the three parameter Gamma distribution are

$$q = \sum_{k=1}^N (U_{(k)} - U_{(1)} + V_{(k)} - V_{(1)} + W_{(k)} - W_{(1)}) \sim \Gamma(3(N-1), \lambda, 0),$$

$$U_{(1)} \sim \Gamma(1, \lambda/N, \phi_p + \tau), V_{(1)} \sim \Gamma(1, \lambda/N, \phi_q + \tau), W_{(1)} \sim \Gamma(1, \lambda/N, \phi_q - \phi_p + \tau).$$

Next, defining $g(\mathbf{T}')$ and $h(\mathbf{T}')$ as functions of \mathbf{T}' such that $E[g(\mathbf{T}')] = E[h(\mathbf{T}')] = \Phi_{\mathbf{S}}$,

$$E[g(\mathbf{T}') - h(\mathbf{T}')] = E[\pi(\mathbf{T}')] = 0 \quad \forall \Phi_{\mathbf{S}}$$

where the expectation is taken with respect to $p(\mathbf{T}'; \Phi_{\mathbf{S}})$. As a result, since the domain of q is also dictated by $U_{(1)}$, $V_{(1)}$ and $W_{(1)}$,

$$\int \int \int \int_{R_{U_{(1)}, V_{(1)}, W_{(1)}}} \pi(q, U_{(1)}, V_{(1)}, W_{(1)}) \cdot \frac{\lambda^{-3(N-1)}}{\Gamma[3(N-1)]} q^{3N-4} e^{-\frac{q}{\lambda}} \cdot \left(\frac{N}{\lambda}\right)^3 e^{-\frac{N}{\lambda}\{U_{(1)}+V_{(1)}+W_{(1)}-2\phi_q-3\tau\}} \cdot dq dU_{(1)} dV_{(1)} dW_{(1)} = 0 \quad \forall \Phi_{\mathbf{S}}$$

where $R_{U_{(1)}, V_{(1)}, W_{(1)}}$ is the region defined by $I[U_{(1)} - \phi_p - \tau]$, $I[V_{(1)} - \phi_q - \tau]$ and $I[W_{(1)} - \phi_q + \phi_p - \tau]$. It follows that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \pi(q, U_{(1)}, V_{(1)}, W_{(1)}) \cdot q^{3N-4} e^{-\frac{q}{\lambda}\{\frac{q}{N}+U_{(1)}+V_{(1)}+W_{(1)}\}} dq dU_{(1)} dV_{(1)} dW_{(1)} = 0 \quad \forall \Phi_{\mathbf{S}}$$

From the uniqueness theorem for the two-sided Laplace transform, it follows that $\pi(\mathbf{T}') = 0$ almost everywhere, resulting in the completeness of \mathbf{T}' , or equivalently \mathbf{T} . Hence, \mathbf{T} is also the minimal sufficient statistics from Bahadur's theorem and the MVUE is the same as $\hat{\Phi}_{\mathbf{S}}$ in (6.7) expressed as

$$\hat{\Phi}_{\mathbf{S}} = \begin{bmatrix} 2V_{(1)} - U_{(1)} - W_{(1)} \\ V_{(1)} - W_{(1)} \\ \frac{1}{3(N-1)} \{3N(U_{(1)} + W_{(1)} - V_{(1)}) + 2(2V_{(1)} - U_{(1)} - W_{(1)}) - (\bar{U} + \bar{V} + \bar{W})\} \\ \frac{N}{3(N-1)} \{(\bar{U} + \bar{V} + \bar{W}) - (U_{(1)} + V_{(1)} + W_{(1)})\} \end{bmatrix},$$

The covariance matrix of this estimator is given by (6.6) and the diagonal elements represent the variance of each individual unknown parameter, whereas the trace of this matrix is the total mean square error or variance for the vector parameter $\Phi_{\mathbf{S}}$.

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

$$\hat{\phi}_q = 2V_{(1)} - U_{(1)} - W_{(1)},$$

and its variance can be written as

$$\text{var}(\phi_q) = \frac{6\lambda^2}{N^2}.$$

3. Minimum Mean Square Error Estimation

Proceeding similarly as before, let $\mathbf{\Delta}_{\text{MMSE}}^{\text{S}} \triangleq [\phi_q \ \phi_p \ \tau]^T$ and $\mathbf{\Xi}_{\text{MMSE}}^{\text{S}} \triangleq \lambda$. If the unique joint minimum variance unbiased estimator is denoted by $[\hat{\mathbf{\Delta}}_{\text{MVUE}}^{\text{S}} \ \hat{\mathbf{\Xi}}_{\text{MVUE}}^{\text{S}}]$, and their covariance matrix from (6.6) is given by

$$\begin{bmatrix} \mathbf{E}_1 & \mathbf{E}_2 \\ \mathbf{E}_3 & \mathbf{E}_4 \end{bmatrix},$$

where

$$\begin{aligned} \mathbf{E}_1 &= \frac{\lambda^2}{N^2} \begin{bmatrix} 6 & 3 & -4 \\ 3 & 2 & -2 \\ -4 & -2 & \frac{9N-8}{3(N-1)} \end{bmatrix}, & \mathbf{E}_2 &= \frac{\lambda^2}{N^2} \begin{bmatrix} 0 \\ 0 \\ -\frac{N}{3(N-1)} \end{bmatrix}. \\ \mathbf{E}_3 &= \frac{\lambda^2}{N^2} \begin{bmatrix} 0 & 0 & -\frac{N}{3(N-1)} \end{bmatrix}, & \mathbf{E}_4 &= \frac{\lambda^2}{3(N-1)}, \end{aligned}$$

then the unique joint minimum mean square error with expected loss invariant under transformations of location and scale is

$$\begin{aligned}\hat{\Delta}_{\text{MMSE}}^{\text{S}} &= \hat{\Delta}_{\text{MVUE}}^{\text{S}} - \frac{1}{\lambda^2} \mathbf{E}_2 \left(1 + \frac{1}{\lambda^2} \mathbf{E}_4 \right)^{-1} \hat{\Xi}_{\text{MVUE}}^{\text{S}}, \\ &= \begin{bmatrix} 2V_{(1)} - U_{(1)} - W_{(1)} \\ V_{(1)} - W_{(1)} \\ \hat{\tau}_{\text{MVUE}}^{\text{S}} + \frac{1}{3(N-1)(3N-2)} \hat{\lambda}_{\text{MVUE}}^{\text{S}} \end{bmatrix}\end{aligned}$$

and

$$\begin{aligned}\hat{\Xi}_{\text{MMSE}}^{\text{S}} &= \left(1 + \frac{1}{\lambda^2} \mathbf{E}_4 \right)^{-1} \hat{\Xi}_{\text{MVUE}}^{\text{S}} \\ &= \frac{N}{3N-2} [(\bar{U} + \bar{V} + \bar{W}) - (U_{(1)} + V_{(1)} + W_{(1)})].\end{aligned}$$

It is evident from above that the MMSE estimator for the desired clock offset parameter of the inactive nodes is the same as the MLE, ordered BLUE and the MVUE for the symmetric exponential link delay model.

C. Summary

In this chapter, three different parameter estimation schemes, namely best linear unbiased estimation using ordered observations, minimum variance unbiased estimation and minimum mean square error estimation with expected loss independent of clock offsets and fixed link delay, are employed to accomplish the goal of accurately synchronizing a wireless sensor network. The results provided here not only apply the above mentioned better schemes to the clock synchronization problem as compared to the MLE derived previously in Chapter V, but also generalize its scope by taking into account asymmetric link delays scenario. In addition, it is shown that under symmetric delay case, the BLUE based on ordered data, the MVUE and the MMSE

estimators coincide with the MLE. These findings are very useful in the realm of wireless sensor networks, where many applications demand tight synchronization among the clocks of the nodes while keeping the spent power at the bare minimum.

CHAPTER VII

CLOCK DRIFT ESTIMATION FOR ACHIEVING LONG-TERM
SYNCHRONIZATION*

Until now, various schemes have been proposed for estimating the clock offset and skew. However, estimating the clock of a node using a linear model is useful only for short-term applications, examples of which are object tracking and surveillance. But it is not sufficient for some applications having stringent and long term clock synchronization requirements such as efficient duty cycling and synchronized sampling because they spend a lot of energy for resynchronization during a given long enough time interval.

To elaborate more on this point, consider the following examples. FTSP has to resynchronize the nodes in the network *every minute* to achieve $90 \mu s$ synchronization error, even though it is the most efficient time synchronization protocol reported thus far and implemented on real testbeds yielding very good results [13]. In addition, the Center for Embedded Networked Sensing (CENS) deployment at James Reserves [45] uses RBS to synchronize the nodes after every 5 minutes and the shooter localization system [44] implements FTSP to synchronize once every 45 seconds. Due to these reasons, even though RBS and FTSP estimate the clock skew alongside clock offset using linear regression, they are insufficient in practice for achieving long term synchronization since they are confined to estimating up to the parameter of first order (clock skew) only. Hence, for achieving this goal of long term synchronization, a bet-

*Part of this chapter is reprinted with permission from “Clock Estimation for Long-Term Synchronization in Wireless Sensor Networks with Exponential Delays” by Qasim M. Chaudhari and Erchin Serpedin, EURASIP Journal on Advances in Signal Processing, Volume 2008, Article ID 219458, 6 pages, 2008.

ter modeling of clock relationship with the reference node is required. In this section, this problem is targeted through extending the linear model between two clocks to a quadratic one and then the clock parameters of clock offset, skew and drift are jointly estimated.

Estimating this second order parameter of the model, called clock drift, is important in light of the reasons mentioned above and finding the MLE is attractive due to its optimal properties for a large number of observations as mentioned before. It should be noted here that although the estimation of clock parameters using a quadratic model is computationally more demanding than using the linear model, it helps in maintaining long-term synchronization among the nodes and subsequently less frequent communications and power saving. Since it has been reported in [31] that the energy required to transmit 1 bit over 100 meters (3 Joules) is equivalent to the energy required to execute *3 millions* of instructions, a trade-off between spending reduced communication energy on the cost of more computational energy through estimating the long-term drift as well as the offset and the skew between clocks of two nodes is an option worth adopting.

A. Problem Formulation

Now for a similar two-way timing cells exchange as before but with addition of the new quadratic terms is given by

$$\begin{aligned} T_{2,r} &= T_{1,r}^2\varphi + T_{1,r}\omega + \phi + \tau + X_r, \\ T_{3,r} &= T_{4,r}^2\varphi + T_{4,r}\omega + \phi - \tau - Y_r, \end{aligned}$$

where X_r and Y_r are modeled as exponential random variables. Based on the above model, the likelihood function can be written as

$$L(\tau, \phi, \omega, \varphi) = \alpha^{-2N} .e^{-\frac{1}{\alpha} \sum_{i=1}^N (T_{4,r}^2 - T_{1,r}^2) \varphi + \sum_{i=1}^N \{(T_{4,r} - T_{1,r}) \omega + (T_{2,r} - T_{3,r}) - 2\tau\}}$$

$$\prod_{i=1}^N I [+T_{2,r} - T_{1,r}^2 \varphi - T_{1,r} \omega - \phi - \tau \geq 0] \prod_{i=1}^N I [-T_{3,r} + T_{4,r}^2 \varphi + T_{4,r} \omega + \phi - \tau \geq 0].$$

We assume that the clocks can neither stop nor run backwards, which implies that the clock skew $\omega \not\leq 0$ and hence always positive. The actual values of practical clock skew is usually close to 1. Finally, for the simplification of derivation, φ has been assumed to be positive. Following a similar procedure mentioned herein, a negative value of φ will result in the same closed form expression of the MLE.

B. The Estimation Procedure

The constraints present in the likelihood function (7.1) can be expressed equivalently as

$$\tau > 0, \quad \varphi > 0, \quad \omega > 0,$$

$$\infty > \phi > -\infty,$$

$$\tau \leq +T_{2,i} - T_{1,i}^2 \varphi - T_{1,i} \omega - \phi, \quad i = 1, \dots, N \quad (7.1)$$

$$\tau \leq -T_{3,j} + T_{4,j}^2 \varphi + T_{4,j} \omega + \phi, \quad j = 1, \dots, N \quad (7.2)$$

These constraints can be viewed as $2N$ 4-D curves due to the four unknowns. The 3-D region where the two sets of N curves in (7.1) and (7.2) intersect each other yields ϕ in terms of ω and φ as

$$2\phi = (T_{2,i} + T_{3,j}) - (T_{1,i}^2 + T_{4,j}^2) \varphi - (T_{1,i} + T_{4,j}) \omega, \quad i, j = 1, \dots, N \quad (7.3)$$

Plugging it back in (7.1), the sets of constraints can now be written as

$$\tau \leq T_{2,i} - T_{1,i}^2 \varphi - T_{1,i} \omega - \frac{1}{2} [(T_{2,i} + T_{3,j}) - (T_{1,i}^2 + T_{4,j}^2) \varphi - (T_{1,i} + T_{4,j}) \omega],$$

$$i, j = 1, \dots, N$$

or equivalently,

$$2\tau \leq (T_{2,i} - T_{3,j}) + (T_{4,j}^2 - T_{1,i}^2) \varphi + (T_{4,j} - T_{1,i}) \omega, \quad i, j = 1, \dots, N \quad (7.4)$$

The above inequalities in (7.4) represent a 3-D region due to three unknowns consisting of N^2 surfaces forming the boundary of the support region. To find this boundary of the support region as a function of φ only, the intersection of these surfaces in (7.4) with each other are

$$\begin{aligned} \omega &= \frac{[(T_{2,k} - T_{3,l}) - (T_{2,i} - T_{3,j})] + [(T_{4,l}^2 - T_{1,k}^2) - (T_{4,j}^2 - T_{1,i}^2)] \varphi}{(T_{4,j} - T_{1,i}) - (T_{4,l} - T_{1,k})}, \\ &= u_a + v_a \varphi, \end{aligned} \quad (7.5)$$

where

$$\begin{aligned} u_a &= \frac{(T_{2,k} - T_{3,l}) - (T_{2,i} - T_{3,j})}{(T_{4,j} - T_{1,i}) - (T_{4,l} - T_{1,k})}, \\ v_a &= \frac{(T_{4,l}^2 - T_{1,k}^2) - (T_{4,j}^2 - T_{1,i}^2)}{(T_{4,j} - T_{1,i}) - (T_{4,l} - T_{1,k})}, \end{aligned}$$

and a is a simplified index notation as a function of the indices (i, j, k, l) . Now plugging (7.5) into (7.4) yields the support region in terms of τ as a function of φ

only as

$$\begin{aligned}
2\tau &\leq (T_{2,m} - T_{3,n}) + (T_{4,n} - T_{1,m}) \frac{(T_{2,p} - T_{3,q}) - (T_{2,m} - T_{3,n})}{(T_{4,n} - T_{1,m}) - (T_{4,q} - T_{1,p})} \\
&+ (T_{4,n}^2 - T_{1,m}^2)\varphi + (T_{4,n} - T_{1,m}) \frac{(T_{4,q}^2 - T_{1,p}^2) - (T_{4,n}^2 - T_{1,m}^2)}{(T_{4,n} - T_{1,m}) - (T_{4,q} - T_{1,p})} \varphi. \\
&= \frac{(T_{4,n} - T_{1,m})(T_{2,p} - T_{3,q}) - (T_{2,m} - T_{3,n})(T_{4,q} - T_{1,p})}{(T_{4,n} - T_{1,m}) - (T_{4,q} - T_{1,p})} \\
&+ \frac{(T_{4,n} - T_{1,m})(T_{4,q}^2 - T_{1,p}^2) - (T_{4,n}^2 - T_{1,m}^2)(T_{4,q} - T_{1,p})}{(T_{4,n} - T_{1,m}) - (T_{4,q} - T_{1,p})} \varphi. \\
&= w_b + z_b \varphi, \tag{7.6}
\end{aligned}$$

where

$$\begin{aligned}
w_b &= \frac{(T_{4,n} - T_{1,m})(T_{2,p} - T_{3,q}) - (T_{2,m} - T_{3,n})(T_{4,q} - T_{1,p})}{(T_{4,n} - T_{1,m}) - (T_{4,q} - T_{1,p})}, \\
z_b &= \frac{(T_{4,n} - T_{1,m})(T_{4,q}^2 - T_{1,p}^2) - (T_{4,n}^2 - T_{1,m}^2)(T_{4,q} - T_{1,p})}{(T_{4,n} - T_{1,m}) - (T_{4,q} - T_{1,p})},
\end{aligned}$$

and b is again a simplified index notation as a function of the indices (m, n, p, q) . Now the final form and uniqueness of the MLE can be proved by the following theorem.

Lemma 6. *The MLE $(\hat{\varphi}, \hat{\tau}, \hat{\omega}, \hat{\phi})$ is unique and is given by that intersection of two curves on the boundary of the support region in (7.6) where the term $\sum_{r=1}^N \{(T_{4,r}^2 - T_{1,r}^2) + v_a(T_{4,r} - T_{1,r})\} - Nz_b$ is negative for one curve and positive for the other.*

Proof. The MLE $(\hat{\varphi}, \hat{\tau}, \hat{\omega}, \hat{\phi})$ can be derived by the following observations.

1. It is clear that the MLE lies on the boundary of the support region. This is because for any τ lying somewhere within the support region, the likelihood function (7.1) can be further increased by increasing τ until it reaches the boundary of the support region.
2. Maximizing the likelihood function in (7.1) is equivalent to minimizing the exponential function argument $\Phi = \sum_{r=1}^N [(T_{4,r}^2 - T_{1,r}^2)\varphi + (T_{4,r} - T_{1,r})\omega + (T_{2,r} - T_{3,r})$

– $2\tau]$ in the likelihood function expression. Therefore, plugging (7.5) and (7.6) into the expression for Φ , it can be written in the form of a set $\phi_{a,b}$ depending on indices a and b as

$$\begin{aligned}\phi_{a,b} &= \sum_{r=1}^N [(T_{4,r}^2 - T_{1,r}^2)\varphi + (T_{4,r} - T_{1,r})(u_a + v_a\varphi) + (T_{2,r} - T_{3,r}) - (w_b + z_b\varphi)], \\ &\propto \left[\sum_{r=1}^N \{(T_{4,r}^2 - T_{1,r}^2) + v_a(T_{4,r} - T_{1,r})\} - Nz_b \right] \varphi.\end{aligned}$$

3. Starting from z_b corresponding to $\min_b\{w_b\}$ and evaluating $\phi_{a,b}$ on each subsequent z_b on the boundary of the support region, observe that for each particular segment, $\phi_{a,b}$ can be minimized by taking the largest possible $\hat{\varphi}$ if the term $\sum_{r=1}^N \{(T_{4,r}^2 - T_{1,r}^2) + v_a(T_{4,r} - T_{1,r})\} - Nz_b$ is negative and by taking the smallest possible $\hat{\varphi}$ if $\sum_{r=1}^N \{(T_{4,r}^2 - T_{1,r}^2) + v_a(T_{4,r} - T_{1,r})\} - Nz_b$ is positive.
4. Since the boundary of the support region is formed by the curves in (7.6) in an order of decreasing slopes $\{z_b\}$, the intersection where the sign of $\sum_{r=1}^N \{(T_{4,r}^2 - T_{1,r}^2) + v_a(T_{4,r} - T_{1,r})\} - Nz_b$ (and hence the sign of $\phi_{a,b}$) changes from negative to positive occurs only once. Therefore, the MLE must be unique.
5. Let $c = \min_a\{v_a\}$ and $s = \{a\} \setminus c$. Now comparing $\phi_{c,b}$ and $\phi_{s,b}$ on the boundary of the support region yields the following three options.
 - The signs of both $\phi_{s,b}$ and $\phi_{c,b}$ change at the same intersection of curves in (7.6). In this case, $\phi_{c,b} < \phi_{s,b}$ since $v_c < v_s$.
 - The sign change for $\phi_{s,b}$ occurs at an intersection of the curves in (7.6) to the right of the intersection where the sign change for $\phi_{c,b}$ occurs. This is not possible because for the same z_b , $\phi_{s,b}$ must have a sign change at or to the left of the intersection where the same occurs for $\phi_{c,b}$.
 - The sign of $\phi_{s,b}$ changes at an intersection of curves in (7.6) which is to the

left of the intersection where the sign change for $\phi_{c,b}$ occurs. Now even on intersection 1, $\phi_{c,b} < \phi_{s,b}$ since $v_c < v_s$. Due to the continuity of $\phi_{c,b}$ (and hence the continuity of the likelihood function) on the support region, $\phi_{c,b}$ can be further decreased by increasing φ until it touches the intersection 2. Therefore, $a = c$ should be used to find the index b corresponding to the minimum of the set $\phi_{c,b}$.

6. Finally, in the light of above observations, by checking the sign of the expression $\sum_{r=1}^N \{(T_{4,r}^2 - T_{1,r}^2) + v_c(T_{4,r} - T_{1,r})\} - Nz_b$ for each b , we conclude that the MLE $\hat{\varphi}$ can be expressed as

$$\hat{\varphi} = \left[\frac{D_{(4,n)(1,m)}D_{(2,p)(3,q)} - D_{(2,m)(3,n)}D_{(4,q)(1,p)}}{D_{(4,n)(1,m)} - D_{(4,q)(1,p)}} - \frac{D_{(4,j)(1,i)}D_{(2,k)(3,l)} - D_{(2,i)(3,j)}D_{(4,l)(1,k)}}{D_{(4,j)(1,i)} - D_{(4,l)(1,k)}} \right] / \left[\frac{D_{(4,j)(1,i)}D_{(4,l)(1,k)}^2 - D_{(4,j)(1,i)}^2D_{(4,l)(1,k)}}{D_{(4,j)(1,i)} - D_{(4,l)(1,k)}} - \frac{D_{(4,n)(1,m)}D_{(4,q)(1,p)}^2 - D_{(4,n)(1,m)}^2D_{(4,q)(1,p)}}{D_{(4,n)(1,m)} - D_{(4,q)(1,p)}} \right], \quad (7.7)$$

where the difference of any two timestamps are denoted by their indices as $D_{(,) (,)}$ (e.g., $T_{4,n} - T_{1,m}$ as $D_{(4,n)(1,m)}$) and the square of timestamps as $D_{(,) (,)}^2$ (e.g., $T_{4,j}^2 - T_{1,i}^2$ as $D_{(4,j)(1,i)}^2$). where the indices $\{i, j, k, l, m, n, p, q\}$ correspond to the two set of curves in (7.6) for which the sign of $\sum_{r=1}^N \{(T_{4,r}^2 - T_{1,r}^2) + v_c(T_{4,r} - T_{1,r})\} - Nz_b$ changes from negative to positive. Consequently, plugging $\hat{\varphi}$ in (7.6), (7.5) and (7.3), we can write $\hat{\tau}, \hat{\omega}, \hat{\phi}$ as

$$\begin{aligned}
\hat{\tau} &= \frac{1}{2} \frac{D_{(4,j)(1,i)} D_{(2,k)(3,l)} - D_{(2,i)(3,j)} D_{(4,l)(1,k)}}{D_{(4,j)(1,i)} - D_{(4,l)(1,k)}} \\
&\quad + \frac{1}{2} \frac{D_{(4,j)(1,i)} D_{(4,l)(1,k)}^2 - D_{(4,j)(1,i)}^2 D_{(4,l)(1,k)}}{D_{(4,j)(1,i)} - D_{(4,l)(1,k)}} \hat{\varphi}, \\
\hat{\omega} &= \frac{D_{(2,k)(3,l)} - D_{(2,i)(3,j)}}{D_{(4,j)(1,i)} - D_{(4,l)(1,k)}} + \frac{\left[D_{(4,l)(1,k)}^2 - D_{(4,j)(1,i)}^2 \right]}{D_{(4,j)(1,i)} - D_{(4,l)(1,k)}} \hat{\varphi}, \\
\hat{\phi} &= \frac{T_{2,i} + T_{3,j} - (T_{1,i}^2 + T_{4,j}^2) \hat{\varphi} - (T_{1,i} + T_{4,j}) \hat{\omega}}{2}.
\end{aligned}$$

□

Algorithm 6 presents in detail the steps required to find this MLE $(\hat{\varphi}, \hat{\tau}, \hat{\omega}, \hat{\phi})$. As N becomes large, clock drift estimation becomes particularly useful for capturing the clock dynamics in a better way. The complete procedure for finding this MLE $(\hat{\varphi}, \hat{\tau}, \hat{\omega}, \hat{\phi})$ is explained in Algorithm 6. Algorithm 6 starts from the curve in (7.6) for which w has the least value. It selects the intersection of this curve with the neighboring curve intersecting it, and it checks the sign change condition of $\sum_{r=1}^N \{(T_{4,r}^2 - T_{1,r}^2) + v_c(T_{4,r} - T_{1,r})\} - Nz_b$. If the condition is not satisfied, the first curve is discarded and the same procedure is repeated for the second curve and so on until the same condition is satisfied.

C. Summary

Using a quadratic model for the relationship between the clocks of two nodes with a two-way timing message exchange mechanism, we have derived the MLE for the clock offset, skew, drift and the fixed delay between the two nodes. In addition, complete steps for the algorithm required to find this MLE are also presented.

Algorithm 6 ML estimation for $\hat{\varphi}$, τ , $\hat{\omega}$, and $\hat{\phi}$

1: Compute the set $\{v_a\}$ and $\{z_b\}$;

2: $c = \min_a \{v_a\}$;

3: $(i, j, k, l) \longrightarrow \min \{w_b\}$;

LABEL:

$$4: \varphi^{m,n,p,q} = \left[\frac{D_{(4,n)(1,m)} D_{(2,p)(3,q)} - D_{(2,m)(3,n)} D_{(4,q)(1,p)}}{D_{(4,n)(1,m)} - D_{(4,q)(1,p)}} - \frac{D_{(4,j)(1,i)} D_{(2,k)(3,l)} - D_{(2,i)(3,j)} D_{(4,l)(1,k)}}{D_{(4,j)(1,i)} - D_{(4,l)(1,k)}} \right] / \left[\frac{D_{(4,j)(1,i)} D_{(4,l)(1,k)}^2 - D_{(4,j)(1,i)}^2 D_{(4,l)(1,k)}}{D_{(4,j)(1,i)} - D_{(4,l)(1,k)}} - \frac{D_{(4,n)(1,m)} D_{(4,q)(1,p)}^2 - D_{(4,n)(1,m)}^2 D_{(4,q)(1,p)}}{D_{(4,n)(1,m)} - D_{(4,q)(1,p)}} \right];$$

5: $(e, f, g, h) = \arg \min_{m,n,p,q} \{\varphi^{m,n,p,q}\}$;

6: **if** $\left[\sum_{r=1}^N \{D_{(4,r)(1,r)}^2 + v_c D_{(4,r)(1,r)}\} - N z_b \right]^{i,j,k,l} \times \left[\sum_{r=1}^N \{D_{(4,r)(1,r)}^2 + v_c D_{(4,r)(1,r)}\} - N z_b \right]^{e,f,g,h} < 0$ **then**

7: $\hat{\varphi} = \varphi^{e,f,g,h}$;

$$\hat{\tau} = \frac{1}{2} \frac{D_{(4,f)(1,e)} D_{(2,g)(3,h)} - D_{(2,e)(3,f)} D_{(4,h)(1,g)}}{D_{(4,f)(1,e)} - D_{(4,h)(1,g)}} + \frac{1}{2} \frac{D_{(4,f)(1,e)} D_{(4,h)(1,g)}^2 - D_{(4,f)(1,e)}^2 D_{(4,h)(1,g)}}{D_{(4,f)(1,e)} - D_{(4,h)(1,g)}} \hat{\varphi},$$

$$\hat{\omega} = \frac{D_{(2,g)(3,h)} - D_{(2,e)(3,f)} + [D_{(4,h)(1,g)}^2 - D_{(4,f)(1,e)}^2]}{D_{(4,f)(1,e)} - D_{(4,h)(1,g)}},$$

$$\hat{\phi} = \frac{1}{2} [T_{2,e} + T_{3,f} - (T_{1,e}^2 + T_{4,f}^2) \hat{\varphi}] - \frac{1}{2} [(T_{1,e} + T_{4,f}) \hat{\omega}];$$

8: **else**

9: Discard (i, j, k, l) curve;

10: $(i, j, k, l) = (e, f, g, h)$;

11: goto LABEL;

12: **end if**

CHAPTER VIII

JOINT SYNCHRONIZATION OF CLOCK OFFSET AND SKEW IN A
RECEIVER-RECEIVER PROTOCOL*

Turning our attention in this chapter towards a general receiver-receiver protocol in which a master node sends reference broadcasts to the neighboring nodes, the joint MLE for clock phase offset and skew under exponential noise model is formulated and found via a direct algorithm. The Gibbs Sampler is also proposed for joint clock phase offset and skew estimation and shown to provide superior performance relative to JMLE. Lower and upper bounds for the mean-square errors (MSE) of JMLE and Gibbs Sampler are introduced in terms of the MSE of the MVUE and the conventional BLUE, respectively.

A. Modeling Assumptions

In a general receiver-receiver protocol, a transmitter node broadcasts N synchronization signals and the receiver nodes put time-stamps on these signals. Then, for efficient implementation, the receivers pass the data consisting of the time-stamps to the transmitter where the clock offsets and skews between different pairs of nodes are calculated. By the help of this protocol, two of the main error sources of clock synchronization are eliminated, which are uncertainties at *Send Time* and *Access Time*. Furthermore, the difference between propagation times is negligible compared to the uncertainty at *Receive Time*, which becomes the only error source. Therefore, the i^{th}

*Part of this chapter is reprinted with permission from “On the Joint Synchronization of Clock Offset and Skew in RBS-Protocol” by Ilkay Sari, Erchin Serpedin, Kyoung-Lae Noh, Qasim M. Chaudhari and Bruce Suter, IEEE Transactions on Communications, Volume 56, Issue 5, May 2008, Page(s): 700 - 703.

time-stamps at the receivers X and Y are given by

$$X[i] = T_1 + \theta_x + \beta_x \tau[i] + v_{x,\lambda_x}[i], \quad Y[i] = T_1 + \theta_y + \beta_y \tau[i] + v_{y,\lambda_y}[i], \quad (8.1)$$

where T_1 stands for the time on the transmitter when it sends the first synchronization signal, θ_x and β_x stand for the offset and skew between the clocks of the receiver X and the transmitter, $\tau[i]$ stands for the difference between T_1 and the time of i^{th} synchronization signal (with respect to the transmitter's clock) and $v_{x,\lambda_x}[i]$ stands for the exponential iid (independently and identically distributed) noise (with mean $1/\lambda_x$), with $i = 1, \dots, N$. The parameters to be estimated, the offset and skew between the clocks of the nodes X and Y , are given by the following equations:

$$\Theta = \theta_x - \theta_y, \quad \beta = \beta_x - \beta_y. \quad (8.2)$$

B. JML Estimation of the Offset and Skew

The estimation of clock skew becomes more important in the context of energy-constrained sensor networks. [46] shows that under uniform noise, there are infinite solutions for ML estimation. Besides, the support of likelihood function is not convex which leaves out the possibility of taking the mean of all equally likely solutions. In this letter, we will consider the case described in (8.1). As long as the two parameter sets $\{\theta_x, \beta_x, \lambda_x\}$ and $\{\theta_y, \beta_y, \lambda_y\}$ do not have a direct relationship and the noise sources in different nodes are independent (both of which are realistic assumptions), we can find the JML-estimator for Θ and β without loss of any information by estimating the parameters (θ_x, β_x) and (θ_y, β_y) separately and plugging these estimates back into (8.2). Thus, we will concentrate on the estimation of θ_x and β_x . First of all, for simplicity, we will assume that $\tau[i] = i - 1$ and $T_1 = 0$, then the likelihood function

becomes

$$\begin{aligned}
 L(\theta_x, \beta_x) &= \prod_{i=1}^N \lambda_x e^{-\lambda_x (X[i] - (\theta_x + (i-1)\beta_x))} I_{(X[i] \geq \theta_x + (i-1)\beta_x)} \\
 &= \lambda_x^N e^{-\lambda_x N(\bar{X} - f)} \prod_{i=1}^N I_{(X[i] \geq f_i)}, \tag{8.3}
 \end{aligned}$$

where $f(\theta_x, \beta_x) = \theta_x + \frac{N-1}{2}\beta_x$, $f_i(\theta_x, \beta_x) = \theta_x + (i-1)\beta_x$, \bar{X} stands for the sample mean of observations $X[i]$ ($i = 1, \dots, N$), and $I_{(x \geq a)}$ denotes the indicator function, being equal to 1 when $x \geq a$ and being 0 elsewhere.

Note that in (8.3), the multiplication of indicator functions defines a convex region \mathbf{S} on the parameter space (θ_x, β_x) , with $\mathbf{S} = \{(\theta_x, \beta_x) : \bigcap_{i=1}^N f_i(\theta_x, \beta_x) \leq X[i]\}$. \mathbf{S} has k vertices $\{s_j\}_{j=1}^k$ and $k+1$ edges ($1 \leq k \leq N-1$). Specifically, the shape of this region and the value of k will strongly depend on the ordering of $X[1], \dots, X[N]$. On this region, we have to maximize the objective function, $f(\theta_x, \beta_x) = \theta_x + \frac{N-1}{2}\beta_x$. Since $0 < \frac{N-1}{2} < N-1$, the support of the solution is guaranteed to be a closed-

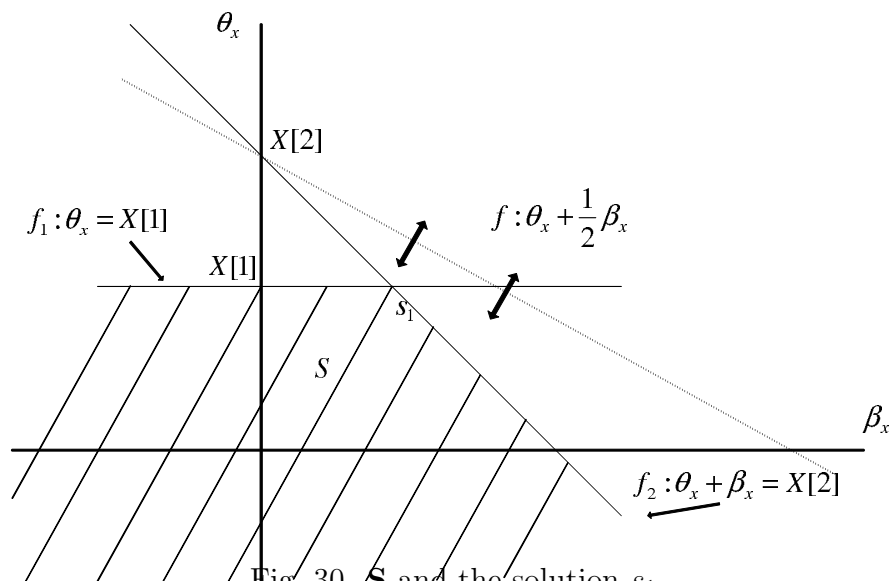


Fig. 30. \mathbf{S} and the solution s_1 .

convex region on the boundary of \mathbf{S} . If $N = 2m$, the solution will be one of the vertices s_j and if $N = 2m - 1$ the solution will assume possibly a segment of the line

$f_m : \theta_x + (m - 1)\beta_x$ (or again one of the vertices s_j , depending on the ordering of the observations). Fig. 30 illustrates these remarks for $N = 2$ and $X[2] > X[1]$. In this illustrative example, since f attains its maximum on s_1 amongst all points on \mathbf{S} , s_1 gives the JML estimation of θ_x and β_x . Before proceeding any further, we have to clarify one more point. In derivations up to now, we assumed that λ_x and λ_y were both known. However, if we assume they are unknown and use the reduced likelihood function for (θ_x, β_x) as in [18], it is straightforward to show that we end up with the same JML solution.

C. Application of Gibbs Sampler

Although it is possible to find the exact solution for the ML-estimate as explained above, we will also apply the Gibbs Sampler to jointly estimate the parameters. Although by using the Gibbs Sampler it is possible to find an approximate JML estimation which is arbitrarily close to the exact one, there are some more important advantages that the Gibbs Sampler will provide us. First of all, it can be shown that the JML estimation $(\hat{\theta}_{x,ML}, \hat{\beta}_{x,ML})$ is biased for finite N . (As an example consider the case in Fig. 30, $E[\hat{\theta}_{x,ML}] = E[X[1]] = E[\theta_x + v_{x,\lambda_x}[1]] = \theta_x + 1/\lambda_x$.) For this reason, we need to look for an MVUE estimator. However, the Neyman-Fisher factorization theorem provides $\min_i((X[i] - \theta_x + (i - 1)\beta_x))$ as sufficient statistics, which is not independent of the parameters to be estimated. On the other hand, if we use the Gibbs Sampler at the end we do not have just a single point estimate but the posterior distribution for the parameters to be estimated as the output. Then, we can either find the JML-estimator or set the corresponding estimator as the mean value of the posterior distribution of the parameter, which will automatically perform the marginalization and will give better results with reduced bias and variance.

Another appealing feature of the Gibbs Sampler is its straightforward extendability for additional unknown parameters. For example, it is possible that λ_x is unknown or in addition to the clock phase offset and skew we could have clock drifts: γ_x and γ_y . Although very important, due to the limited space we did not consider such a scenario in this letter. The drifts will be observed on RHS (right-hand side) of (8.1) as additional terms: $\tau^2[i]\gamma_x$ and $\tau^2[i]\gamma_y$. Definitely, it is straightforward to adapt the Gibbs Sampler to these scenarios.

Before applying it, we will briefly give some information about the Gibbs Sampling. Assume that we have the data vector \mathbf{z} and we want to estimate some parameters $\Phi = [\phi_1, \phi_2, \dots, \phi_M]^T$. For any kind of statistical inference we want to use the joint posterior distribution of the parameters $p(\Phi|\mathbf{z}) \propto p(\mathbf{z}|\Phi)p(\Phi)$ (in point estimation, prior distribution $p(\Phi)$ is chosen as noninformative). When it is hard to carry out mathematical derivations on the posterior, we stick to Monte-Carlo methods, i.e., to draw as many samples as possible from the posterior so that the inference we make using these samples will be arbitrarily close to the exact solution. When it is hard to draw from the joint posterior directly, MCMC (Markov Chain Monte-Carlo) type of iterative methods will be used. That resumes to setting up a Markov chain whose stationary distribution is the joint posterior we need. One convenient way to do this is to use Gibbs Sampling in which we iteratively draw samples from one-dimensional conditionals $p(\phi_i|\mathbf{z}, \overline{\Phi}_i)$, where $\overline{\Phi}_i$ is an $(M - 1) \times 1$ vector with entries $\{\phi_j\}_{j \neq i}$. Under mild conditions, these one dimensional conditional distributions uniquely determine the joint posterior distribution [47].

Specifically, the general algorithm for Gibbs Sampling with initial values $\Phi^{(0)} = [\phi_1^{(0)}, \dots, \phi_M^{(0)}]$ is to iterate the following:

- Draw $\phi_1^{(1)}$ from $p(\phi_1|\mathbf{z}, \phi_2^{(0)}, \dots, \phi_M^{(0)})$

- Draw $\phi_2^{(1)}$ from $p(\phi_2|\mathbf{z}, \phi_1^{(1)}, \phi_3^{(0)}, \dots, \phi_M^{(0)})$
- \vdots
- Draw $\phi_M^{(1)}$ from $p(\phi_M|\mathbf{z}, \phi_1^{(1)}, \dots, \phi_{M-1}^{(1)})$.

After a threshold value t , the set $\{\Phi^{(t)}, \Phi^{(t+1)}, \dots\}$ behaves like samples from the joint posterior of the parameters.

One important point is that the joint posterior distribution should be proper. Otherwise the Gibbs Sampler always converges to some local points, but not necessarily to a meaningful one [48]. For this reason to assure that the posterior is proper, in application of Gibbs Sampler to the point estimation, priors are not directly chosen as flat, but they are chosen from conjugate families and then their parameters arranged so as to have noninformative priors. However in our case, the likelihood function itself can be used as posterior distribution, since its integral is always bounded and positive-valued which makes it proper. We do not need to use any other type of priors but flat. Then in our case, using (8.3), the procedure becomes

- Draw $\theta_x^{(1)}$ from $\propto e^{\lambda_x N \theta_x} I(\theta_x \leq \min_i (X[i] - (i-1)\beta_x^{(0)}))$
- Draw $\beta_x^{(1)}$ from $\propto e^{\lambda_x \frac{N(N-1)}{2} \beta_x} I(\beta_x \leq \min_i (\frac{X[i] - \theta_x^{(1)}}{i-1}))$.

For θ_x^{t+1} , we will draw a sample from the exponential distribution with parameter $\lambda_x N$, multiply it with -1 and add $\min_i (X[i] - (i-1)\beta_x^{(t)})$ to it. The procedure for β_x^{t+1} is similar. Note that if λ_x were unknown, we would utilize the Gamma distribution to draw for $\lambda_x^{(t+1)}$.

D. Performance Bounds and Simulations

In this part, we will look at the performances of the Gibbs Sampler and the JML-estimator. However, it will be useful to have some benchmarks with whom to compare

their performances. First we will look for lower bounds. Since the likelihood function does not satisfy the regularity conditions required by CRLB (Cramer-Rao Lower Bound), calculating CRLBs is dropped out of the list. One possible lower bound can be found by assuming that all the parameters are known but the one to be bounded, which reduces the problem to the well-known derivation of the bound for a single unknown parameter in exponential noise. Then we can find the MVUE both for the phase offset and skew in closed forms. For θ_x , it is derived that the MVUE becomes

$$\hat{\theta}_{x,MVUE} = \min_i (X[i] - (i-1)\beta_x) - \frac{1}{N\lambda_x}, \quad (8.4)$$

and the MSE of the estimator equals $1/(N\lambda_x)^2$. For β_x , the likelihood function is

$$L(\beta_x) = C e^{\lambda_x \frac{N(N-1)}{2} \beta_x} \prod_{i=2}^N I(\beta_x \leq \frac{X[i] - \theta_x}{i-1}). \quad (8.5)$$

By Factorization Theorem, $\min_i (\frac{X[i] - \theta_x}{i-1})$ is sufficient statistics and it is straightforward to show that it is also complete. This result can be established by following the similar lines of proof as it is done in [49] for θ_x . Then, by Lehmann-Scheffe Theorem, the MVUE for the skew when the offset and λ_x are known takes the form:

$$\hat{\beta}_{x,MVUE} = \min_i (\frac{X[i] - \theta_x}{i-1}) - \frac{2}{\lambda_x N(N-1)}. \quad (8.6)$$

The MSE of the estimator (8.6) is equal to the variance of $Z = \min_{i=2,3,\dots,N} (\frac{X[i] - \theta_x}{i-1})$. Thus, we first need to determine the distribution of Z . From the theory of order statistics, the distribution of the minimum of a sample set is given by $F(z) = 1 - (1 - F_2(z))(1 - F_3(z)) \dots (1 - F_N(z))$, where $F_i(z) = Pr(\frac{X[i] - \theta_x}{i-1} \leq z) = Pr(v_{x,\lambda_x}[i] \leq (i-1)(z - \beta_x)) = (1 - e^{\lambda_x(i-1)(z-\beta_x)})I(z \geq \beta_x)$. Then the distribution becomes

$$F(z) = 1 - e^{\lambda_x(z-\beta_x)(1+2+\dots+N-1)} = 1 - e^{\lambda_x(z-\beta_x)\frac{N(N-1)}{2}}, \quad (8.7)$$

which is an exponential distribution with the scale parameter $\frac{\lambda_x N(N-1)}{2}$ and the location parameter β_x . The MSE of $\hat{\beta}_{x,MVUE}$ equals the variance of Z which is $4/(\lambda_x N(N-1))^2$. Therefore, we do not expect the MSE of joint estimator for (θ_x, β_x) to decay faster than $\propto (1/N^2, 1/N^4)$.

We will also consider the BLUE, since it will represent an upper bound. Here, the same notation is used as [46] except that \mathbf{X} is replaced with \mathbf{A} ($\mathbf{A} \triangleq [\mathbf{1}, \mathbf{x}]$, where $\mathbf{1} = [1, 1, \dots, 1]^T$ and $\mathbf{x} = [0, 1, \dots, N-1]^T$) to prevent possible confusion. Since noise is not zero-mean in our model unlike [46], we need to subtract $1/\lambda_x$ from the resulting linear estimate of θ_x so as to end up with the BLUE. Then we have $[\hat{\theta}_{x,BLUE}, \hat{\beta}_{x,BLUE}]^T = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{X} - [\frac{1}{\lambda_x}, 0]^T$. It is known that $\text{var}([\hat{\theta}_{x,BLUE}, \hat{\beta}_{x,BLUE}]^T) = \text{diag}\{1/\lambda_x^2 (\mathbf{A}^T \mathbf{A})^{-1}\} \propto [1/N, 1/N^3]^T$. The MSE of the

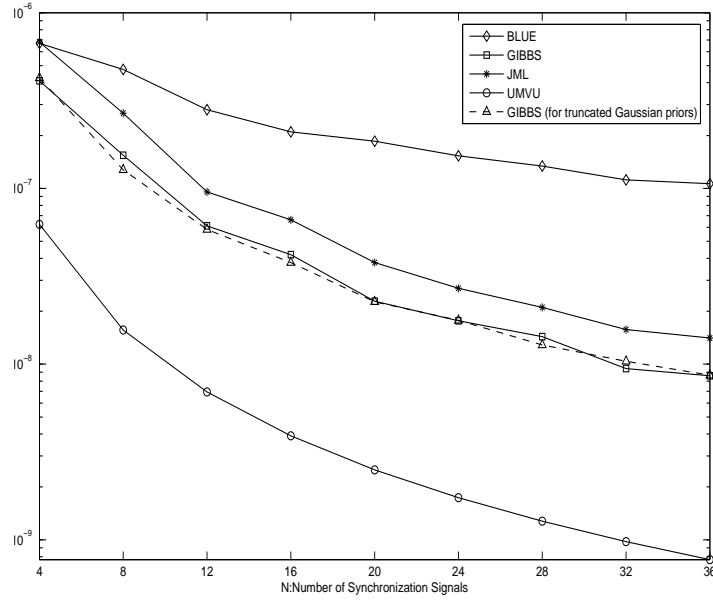


Fig. 31. MSE for $\hat{\theta}_{x,BLUE}$, $\hat{\theta}_{x,JML}$, $\hat{\theta}_{x,GIBBS}$, and $\hat{\theta}_{x,MVUE}$.

Gibbs Sampler and the JML-estimator for $\theta_x = 1$ and $\beta_x = 0.01$ with $\lambda_x = 10^3$ (which makes $\text{var}(v_{x,\lambda_x}) = 10^{-6}$), are presented in Figs. 31 and 32, respectively. In these sim-

ulations, the initial values of clock parameters are chosen as zeros. These figures also

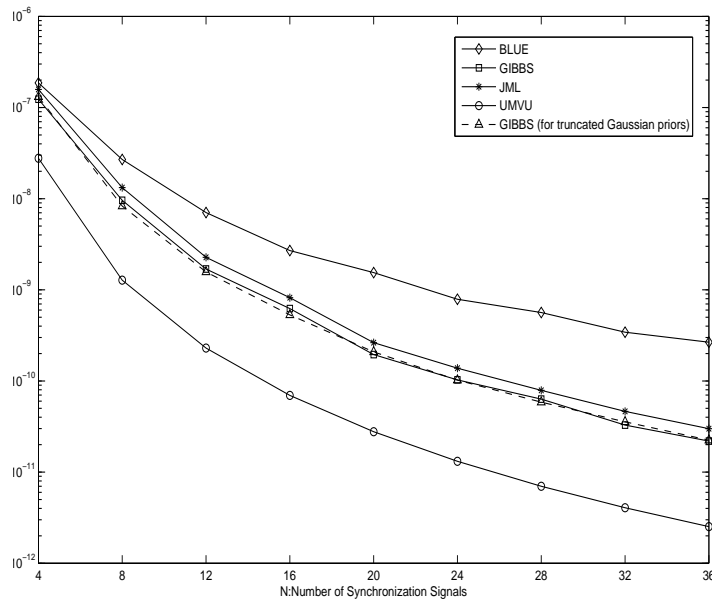


Fig. 32. MSE for $\hat{\beta}_{x,BLUE}$, $\hat{\beta}_{x,JML}$, $\hat{\beta}_{x,GIBBS}$ and $\hat{\beta}_{x,MVUE}$.

include the lower and upper bounds presented above. The MSE are plotted against the number of synchronization signals from 4 to 36. It is interesting to note that the MSE of the Gibbs Sampler and the JML-estimator behave like the lower bound, i.e., decay rates on the order of $\propto 1/N^2$ and $\propto 1/N^4$, respectively. Note also that the Gibbs Sampler performs better with MSE-values around 40% for θ_x and 25% for β_x compared to the corresponding values of JML-estimator. We should also note that the convergence of the Gibbs Sampler is achieved after a number of iterations on the order of N . To shed some light on the sensitivity of the Gibbs Sampler to the prior mismatch, we have also provided some simulation results for the mismatched prior knowledge. This is important for engineers and system designers in order to make proper choice of estimator for their considered systems. Fig. 31 and Fig. 32 show the performance of Gibbs Sampler where we have modeled actual prior as a

truncated Gaussian while the assumed prior in the Gibbs Sampler is uniform. For prior of offset, truncation points have been chosen as 0 and 10 whereas the mean and Standard Deviation of parent Gaussian distribution as 5 and 2 respectively. And for prior of skew, truncation points are chosen as 0 and 1 whereas the mean and Standard Deviation of parent Gaussian distribution as 0.5 and 0.25 respectively.

One drawback of the Gibbs Sampler is definitely its computational complexity. The computational complexity of Gibbs Sampler is affected by the random number generations in each iteration and the number of iterations necessary to converge. Although the Gibbs Sampler clearly requires more computations, the required level of precision can be achieved by lesser number of signal transmissions. Hence, there is a tradeoff between the complexity and the gains achieved by Gibbs Sampler.

E. Summary

Under the exponential noise model, it has been proved that the JMLE of the skew and the phase offset exists and is either unique or a line segment depending on the magnitudes of the observed data samples. At worst, the support of all equally likely solutions is a closed-convex set (a line segment). The setting was convenient to apply Gibbs Sampler which further increased the performance of JML-estimator. The performances of both estimators (JML and Gibbs Sampler) scale with the same power-law (with respect to the number of synchronization signals N). Lower and upper-bounds for the performance of JML and Gibbs Sampler estimators were also presented in terms of the MSE-performances of MVUE estimator and BLUE, respectively.

CHAPTER IX

CONCLUSIONS AND FUTURE WORK

Clock synchronization among the nodes in WSNs is a challenging task due to conflicting requirements of increased accuracy and restricted power consumption. This dissertation targets the clock synchronization problem in a general sender-receiver and receiver-receiver timing packet 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 minimum variance unbiased estimate is also obtained, and compared with the maximum likelihood estimator to find the superior estimator at various regions of asymmetry. Then the MLEs of both the clock offset and skew for any general time synchronization protocol involving a two way message exchange mechanism are derived assuming both Gaussian and exponential delays and the algorithms used for finding these MLEs are also detailed. With sacrificing some performance, two reduced complexity schemes are proposed to estimate the clock parameters in an economical way. The maximum likelihood, best linear unbiased, minimum variance unbiased and minimum mean square error estimators are then derived for the inactive nodes overhearing a timing exchange between a pair of nodes. Afterwards, a procedure to estimate the clock drift is also presented for increased accuracy and longer synchronization period. Finally, the joint maximum likelihood estimator is derived in a general receiver-receiver setting and Gibbs Sampler is applied to increase the accuracy.

There are several future research directions that can be investigated. First, re-examination of this problem utilizing modern statistical inference techniques such as bootstrap and jackknife is an motivating area worth exploring. The methodology adopted in this dissertation can be extended to analyze other time synchro-

nization protocols for both single hop and multihop cases. In addition, finding the CRLBs for the clock offset and skew estimators derived here represents an important open research problem. However, it must be noted that the CRLB in this case can not be derived by the same procedure as in Chapter III. The reason is that $\min_{1 \leq k \leq N} (T_{2,k}/\theta_B - T_{1,k})$ and $\min_{1 \leq k \leq N} (T_{4,k} - T_{3,k}/\theta_B)$ in Case II intersecting at optimal $(\hat{d}, \hat{\theta}_B)$ (and similar curves in other cases) do not correspond to $\min_{1 \leq k \leq N} X_k$ and $\min_{1 \leq k \leq N} Y_k$ respectively. In addition, all the order statistics from an exponential distribution, except the first, do *not* have exponential distribution. Additionally, exploring the effects of violation of iid assumption for the random delays, missing data points due to communication losses, or quantization errors are interesting open problems.

This work might also formulate a very good justification for adopting a Bayesian estimation framework where the currently inactive nodes observe a two-way timing cell exchange mechanism as usual, and adopting a Bayesian approach for estimating their clock offset using the current estimator PDF as prior might considerably improve the synchronization accuracy. Moreover, by modeling the relationship between two clocks with the addition of higher order terms such as clock skew, the parameters for long term synchronization can be estimated and used to increase the lifetime of the network. Lastly, formulating a procedure through which the timing error accumulation over a series of hops encompassing the whole network could be quantified is an idea worth exploring.

REFERENCES

- [1] D. Mills, "Internet time synchronization: The network time protocol," *IEEE Transactions on Communications*, Vol. 39, No. 10, pp. 1482-1493, Oct 1991.
- [2] F. Cristian, "Probabilistic clock synchronization," *Distributed Computing* 3, pp. 146-158, Springer-Verlag, 1989.
- [3] K. Arvind, "Probabilistic clock synchronization in distributed systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 5, no. 5, pp. 474-487, May 1994.
- [4] K. Romer, "Time synchronization in ad hoc networks," in *Proceedings of ACM Symposium on Mobile Ad Hoc Networking and Computing*, pp. 173-182, Long Beach, CA, Oct 2001.
- [5] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393-422, Mar 2002.
- [6] B. M. Sadler, "Fundamentals of energy-constrained sensor network systems," *IEEE Aerosp. Elect. Syst. Magazine (Tutorials II)*, vol. 20, no. 8, pp. 17-35, Aug. 2005.
- [7] B. M. Sadler, "Critical issues in energy-constrained sensor networks: synchronization, scheduling, and acquisition," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Philadelphia, PA, Mar 2005.
- [8] B. M. Sadler and A. Swami, "Synchronization in sensor networks: an overview", in *Proceedings of IEEE Military Communications Conference*, Washington DC, Oct 2006.

- [9] B. Sundararaman, U. Buy and A. D. Kshemkalyani, "Clock synchronization for wireless sensor networks: a survey," *Ad-Hoc Networks*, vol. 3, no. 3, pp. 281-323, Mar 2005.
- [10] J. Elson, L. Girod and D. Estrin, "Fine-grained network time synchronization using reference broadcasts," in *Proceedings of the 5th Symposium on Operating System Design and Implementation*, Boston, MA, Dec 2002.
- [11] S. Ganeriwal, R. Kumar and M.B. Srivastava, "Timing synch protocol for sensor networks," In *Proceedings of 1st International Conference on Embedded Network Sensor Systems*, Los Angeles, CA, Nov 2003.
- [12] A. Syed, J. Heidemann, "Time synchronization for high latency acoustic networks," *Proceedings of INFOCOM 2006*, Barcelona, Spain, April, 2006.
- [13] M. Maroti, B. Kusy, G. Simon and A. Ledeczi, "The flooding time synchronization protocol," In *Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems*, ACM Press, pp. 39-49, Baltimore, MD, Nov 2004.
- [14] TinyOS, <http://webs.cs.berkeley.edu/tos/>.
- [15] Kopetz, H., and Schwabl, W., "Global time in distributed real-time systems," *15/89 Technische Universitat Wien*, 1989.
- [16] Horauer, M. Schmid, U., Schossmaier, K., Holler, R., and Kero, N., "Psynutcevaluation of a high precision time synchronization prototype system for Ethernet LANs," *34th Annual Precise Time and Time Interval Meeting (PTTI)*, (December 2002).
- [17] H. S. Abdel-Ghaffar, "Analysis of synchronization algorithm with time-out control over networks with exponentially symmetric delays," *IEEE Transactions on*

- Communications*, Vol. 50, pp. 1652-1661, Oct 2002.
- [18] D. R. Jeske, "On maximum likelihood estimation of clock offset," *IEEE Transactions on Communications*, Vol. 53, No. 1, pp. 53-54, Jan 2005.
- [19] V. Paxson, "On calibrating measurements of packet transit times," *In Proceedings of 7th ACM Sigmetrics Conference*, Vol. 26, pp. 11-21, Jun 1998.
- [20] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 3rd ed., McGraw-Hill, 1991.
- [21] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*, 2nd ed., Addison-Wesley, 1993.
- [22] C. Bovy, H. Mertodimedjo, G. Hooghiemstra, H. Uijterwaal and P. Miegheem, "Analysis of end-to-end delay measurements in Internet," presented at *the Passive and Active Measurements Workshop*, Fort Collins, CO, Mar 2002.
- [23] S. Moon, P. Skelley and D. Towsley, "Estimation and removal of clock skew from network delay measurements," *In Proceedings of the IEEE INFOCOM Conference on Computer Communications*, New York, NY, Mar 1999.
- [24] Hans Eriksson, "MBONE: the multicast backbone," *Communications of ACM*, Vol. 37, No. 8, Aug 1994.
- [25] 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.
- [26] N. Johnson, S. Kotz, N. Balakrishnan, *Continuous Univariate Distributions, Vol. 1*, Wiley, New York, 2nd Edition, 1994.

- [27] S. M. Kay, *Fundamentals of Statistical Signal Processing, Vol. I. Estimation Theory*, Prentice Hall, 1993.
- [28] R. R. Bahadur, "Sufficiency and statistical decision functions," *The Annals of Mathematical Statistics*, Vol. 25, No. 3., Sep 1954.
- [29] S. Dixit and P. Skelly, "MPEG-2 over ATM or video dialtone networks: issues and strategies," *IEEE Network*, Special Issue on Digital Interactive Broadband Video Dial Tone Networks, Vol. 9, No. 5, pp. 30-40, Sep-Oct 1995.
- [30] R. Ramjee, J. Kurose, D. Towsley, and H. Schulzrinne, "Adaptive playout mechanisms for packetized audio applications in wide-area networks," *In Proceeding of INFOCOM*, Toronto, Canada, Apr 1994.
- [31] G. Pottie and W. Kaiser, "Wireless integrated network sensors," *Comm. of the ACM*, vol. 43, no. 5, pp. 51-58, May 2000.
- [32] O. Simeone and U. Spagnolini, "Distributed time synchronization in wireless sensor networks with coupled discrete-time oscillators," *EURASIP Journal on Wireless Communications and Networking*, Vol. 2007, Article ID 57054, 2007.
- [33] N. Khajehnouri and A. H. Sayed, "A distributed broadcasting time-synchronization scheme for wireless sensor networks," *In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Vol. 5, Philadelphia, PA, March 2005.
- [34] Y.-W. Hong and A. Scaglione, "A scalable synchronization protocol for large scale sensor networks and its applications," *IEEE Journal on Selected Areas in Communications*, Vol. 23, No. 5, 2005.

- [35] Q. Li and D. Rus, "Global clock synchronization in sensor networks," *In Proceedings of IEEE Conf. Computer Communications (INFOCOM)*, Vol. 1, Hong Kong, China, Mar 2004.
- [36] S. Ping, "Delay measurement time synchronization for wireless sensor networks," *Intel Research*, IRB-TR-03-013, Jun 2003.
- [37] M. Sichitiu and C. Veerarittiphan, "Simple, accurate Time synchronization for wireless sensor networks," *In Proceedings of IEEE Wireless Communications and Networking Conference (WCNC)*, 2003.
- [38] M. Mock, R. Frings, E. Nett, and S. Trikaliotis, "Continuous Clock Synchronization in Wireless Real-time Applications," *In Proceedings of 19th IEEE Symposium on Reliable Distributed Systems (SRDS)*, Oct 2000.
- [39] S. PalChaudhuri, A. Saha, and D. B. Johnson, "Adaptive clock synchronization in sensor networks" *In Proceeding of the Information Processing in Sensor Networks (IPSN)*, Berkeley, CA, Apr 2004.
- [40] W. Su and I. F. Akyildiz, "Time-diffusion synchronization protocol for wireless sensor networks," *IEEE/ACM Transactions on Networking*, Vol. 13, Apr 2005.
- [41] K.-L. Noh, E. Serpedin and K. Qaraqe, "A new approach for time synchronization in wireless sensor networks: pairwise broadcast synchronization," *IEEE Transactions on Wireless Communications*, accepted for publication.
- [42] 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 Transactions on Communications*, Vol. 55, No. 4, Apr 2007.

- [43] Nancy R. Mann, "Optimum estimators for linear functions of location and scale parameters," *The Annals of Mathematical Statistics*, Vol. 40, No. 6. Dec. 1969.
- [44] G. Simon, M. Maroti, A. Ledeczi, G. Balogh, B. Kusy, A. Nadas, G. Pap, J. Sallai, K. Frampton, "Sensor network- based counter sniper system," *In Proceedings of the Second ACM Conference on Embedded Networked Sensor Systems (SenSys)*, Nov. 2004.
- [45] CENS habitat sensig group at James Reserve. <http://www.jamesreserve.edu/>
- [46] B. M. Sadler, "Local and broadcast clock sync. in a sensor network," *IEEE Sig. Proc. Letters*, vol. 13, pp. 9-12, Jan. 2006.
- [47] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *Jour. Roy. Stat. Soc. Series B (Methodological)*, vol. 36, pp. 192-236, 1974.
- [48] J. P. Hobert and G. Casella, "Functional compatibility, Markov Chains, and Gibbs Sampling with improper posteriors," *Jour. Comp. Graph. Stat.*, vol. 7, pp. 42-60, 1998.
- [49] E. L. Lehmann and G. Casella, *Theory of Point Estimation*, Springer, 1998.

VITA

Qasim Mahmood Chaudhari received the B.E. degree in Electrical Engineering from National University of Sciences and Technology, Rawalpindi, Pakistan, in November 2001 and the M.S. degree in Electrical Engineering from the University of Southern California, Los Angeles, in May 2004. He completed his Ph.D. degree in Electrical Engineering at Texas A&M University, College Station in August 2008.

He worked with the SoC Tools group of Communications Enabling Technologies, Islamabad, Pakistan, from 2001 to 2002 and with the HSDPA performance test team of Qualcomm Inc., San Diego, CA. His research interests lie in the areas of detection and estimation theory, statistical signal processing and clock synchronization in wireless sensor networks. He can be reached through the Department of Electrical and Computer Engineering, Mailstop 3128, Texas A&M University, College Station, TX 77843-3128.