Optimal Training Sequences for Joint Timing Synchronization and Channel Estimation in Distributed Communication Networks

Ali A. Nasir, *Student Member, IEEE*, Hani Mehrpouyan, *Member, IEEE*, Salman Durrani, *Senior Member, IEEE*, Steven D. Blostein, *Senior Member, IEEE*, Rodney A. Kennedy, *Fellow, IEEE*, and Björn Ottersten, *Fellow, IEEE*

***Abstract*—For distributed multi-user and multi-relay cooperative networks, the received signal may be affected by multiple timing offsets (MTOs) and multiple channels that need to be jointly estimated for successful decoding at the receiver. This paper addresses the design of optimal training sequences for efficient estimation of MTOs and multiple channel parameters. A new *hybrid Cramér-Rao lower bound* (HCRB) for joint estimation of MTOs and channels is derived. Subsequently, by minimizing the derived HCRB as a function of training sequences, three training sequence design guidelines are derived and according to these guidelines, two training sequences are proposed. In order to show that the proposed design guidelines also improve estimation accuracy, the *conditional Cramér-Rao lower bound* (ECRB), which is a tighter lower bound on the estimation accuracy compared to the HCRB, is also derived. Numerical results show that the proposed training sequence design guidelines not only lower the HCRB, but they also lower the ECRB and the mean-square error of the proposed maximum a posteriori estimator. Moreover, extensive simulations demonstrate that application of the proposed training sequences significantly lowers the bit-error rate performance of multi-relay cooperative networks when compared to training sequences that violate these design guidelines.**

***Index Terms*—Training sequence (TS), multiple timing offsets (MTOs), synchronization, channel estimation, hybrid Cramér-Rao lower bound (HCRB), distributed communication network, and maximum-a-posteriori (MAP) estimation.**

I. Introduction

A. Motivation and Literature Survey

TRAINING sequences (TSs) are widely used in communication systems for channel estimation and synchronization [1]. Since the choice of TS significantly affects estimation and system performance [2], TS design is an important topic in the field of communications. It is also known that in distributed multi-user and multi-relay cooperative networks, the receiver needs to detect and decode the signals from multiple nodes that are affected by multiple channels, multiple timing offsets (MTOs), and multiple carrier frequency offsets (MCFOs) [3], [4]. However, the main focus in the literature has been on the design of TSs for channel estimation [5]–[18], or MCFO and channel estimation [19]–[23], while relatively little attention has been paid to the topic of TS design for joint MTO and multiple channel estimation [24].

To date, different approaches have been adopted to design TSs that improve estimation performance. In [14], [15], [19], the authors consider specific estimation algorithms and attempt to find TSs that minimize the corresponding variance of the estimation error. However, the resulting TSs are only optimal for the estimation method under consideration and may not have broader applications. In [5], [23], and [24], new TSs that minimize the Cramér-Rao lower bound (CRB), which is the lower bound on the variance of the estimation error of any unbiased estimator, are derived. However, since the CRBs in [5], [23], and [24] are functions of specific channel realizations, the resulting TSs cannot be guaranteed to be optimal for all instances of random channels. In this regard, it is desirable to derive optimal TSs that are independent of the channel realizations.

Even though estimators for obtaining MTOs and channel parameters in multi-user distributed networks, decode-and-forward (DF) cooperative systems, and amplify-and-forward (AF) cooperative systems are proposed in [3], [25], and [26], respectively, the design of optimal TSs is left as an open research area [25], [26]. For the case of *single* timing offset estimation in point-to-point MIMO systems, [27] proposes optimal TSs based on the design in [17]. However, in [2], it is shown that the proposed TSs in [27] are not optimal in general and instead, Walsh sequences are proposed as a more appropriate alternative. More importantly, the solutions in [27] and [2] are not optimal and are only applicable to point-to-point MIMO systems, since they assume that the received signal is affected by a single timing offset. Thus, as shown in this paper, they are not sufficient for the design of optimal TSs for joint channel and MTO estimation in distributed multi-user and multi-relay cooperative networks. Finally, even though, [24] presents some guidelines for the design of optimal TSs for MTO and multiple channel estimation in cooperative multi-relay networks, they are mainly obtained using numerical

Manuscript received July 27, 2012; revised February 9, 2013. The editor coordinating the review of this paper and approving it for publication was C. Tepedelenlioglu.

Ali A. Nasir, Salman Durrani, and Rodney Kennedy are with Research School of Engineering, the Australian National University, Australia (e-mail: {ali.nasir, salman.durrani, rodney.kennedy}@anu.edu.au).

Hani Mehrpouyan is with the Department of Electrical and Computer Engineering, California Sate University, Bakersfield (e-mail: hani.mehr@ieee.org).

Björn Ottersten is with the Department of Interdisciplinary Center for Security, Reliability and Trust (SnT) at the University of Luxembourg, Luxembourg (e-mail: bjorn.ottersten@uni.lu).

Steven D. Blostein is with the Department of Electrical and Computer Engineering, Queen's University, Canada (e-mail: steven.blostein@queensu.ca).

This research was supported under Australian Research Council's Discovery Projects funding scheme (project number DP110102548). Also, the work was partially supported by Natural Research and Engineering Research Council of Canada (Discovery Grant 41731).

Digital Object Identifier 10.1109/TCOMM.2013.053013.120541

investigations, are not based on an analytical framework, and are not as complete as the guidelines provided in this paper.

B. Contributions

In this paper, we address the problem of designing optimal TSs for joint estimation of MTOs and channels in distributed multi-user or cooperative networks. The *hybrid Cramér-Rao lower bound* (HCRB) for joint estimation of deterministic MTOs and random channel parameters is derived. By minimizing the HCRB with respect to the TSs, three criteria for the design of optimal TSs are proposed. The use of the HCRB results in new guidelines with broader applications since, unlike the CRB, the HCRB is not dependant on the specific channel realizations. Moreover, in order to show that the proposed TSs improve MTO and channel estimation accuracy, the *conditional Cramér-Rao lower bound* (ECRB), which is a tighter bound on the mean-square error (MSE) of estimators, and a new maximum a posteriori (MAP) estimator are derived. The main contributions of this paper can be summarized as follows:

- New HCRBs, ECRBs, and a MAP estimator for joint estimation of multiple channels and MTOs are derived. The derivation of the ECRB is motivated by the fact that it serves as a tight lower bound on the estimation accuracy of the MAP estimator.
- By minimizing the derived HCRB for joint estimation of multiple channels and MTOs, three criteria for the design of optimal TSs are formulated. It is shown that the proposed guidelines not only lower the HCRB, but also lower the ECRB and the MSE of the MAP estimator.
- Based on the proposed TS design criteria, two TSs are proposed and it is demonstrated that these TSs result in a lower MSE for joint estimation of MTO and channel parameters when compared to TSs that violate the proposed guidelines. Moreover, simulations show that application of the proposed TSs in a multi-relay cooperative network significantly enhances the network's BER performance.

C. Organization

The remainder of the paper is organized as follows. Section II outlines the system model and the general set of assumptions. In Section III, the HCRB and ECRB for joint estimation of MTOs and channel parameters are derived. In Section IV, the guidelines for designing optimal TSs are proposed while in Section V, the MAP estimator for joint estimation of MTOs and channels are derived. In Section VI, simulation results are presented and Section VII concludes the paper.

D. Notation

Superscripts $(\cdot)^*$, $(\cdot)^H$, and $(\cdot)^T$ denote the conjugate, the conjugate transpose, and the transpose operators, respectively. Bold face lowercase letters, $\mathbf{x}$, bold face uppercase letters, $\mathbf{X}$, are used for vectors and matrices, respectively. $[\mathbf{X}]_{x,y}$ represents the entry in row x and column y of $\mathbf{X}$. $\mathbf{I}_{X\times X}$ and $\mathbf{0}_{X\times X}$ denote the $X \times X$ identity and all zeros matrices, respectively. $|\cdot|$ is the absolute value operator, $|\mathbf{x}|$ and $\|\mathbf{x}\|$ denote the element-wise absolute value and the ℓ_2 norm of $\mathbf{x}$, respectively. $\det(\mathbf{X})$ and $\mathrm{Tr}(\mathbf{X})$ represent the determinant and trace of $\mathbf{X}$, respectively. $\mathrm{diag}(\mathbf{x})$ is used to denote a diagonal matrix, where its diagonal elements are given by $\mathbf{x}$ and $\mathrm{diag}(\mathbf{X})$ is a vector used to denote the diagonal elements of $\mathbf{X}$. $\mathbb{E}[\cdot]$ denotes the expected value of its argument. $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote the real and imaginary parts of a complex quantity, respectively. $\widetilde{x}$ and $\hat{x}$ denote the approximated and estimated value of x, respectively. $\mathcal{N}(\mu,\sigma^2)$ and $\mathcal{CN}(\mu,\sigma^2)$ are used to denote real and complex Gaussian distributions with mean μ and variance σ^2, respectively. $\mathcal{U}(-x,x)$ denotes a uniform distribution between $-x$ and x. $\Delta^{\mathbf{x}}_{\mathbf{x}} f(\cdot) \triangleq \frac{\partial}{\partial \mathbf{x}}\left[\frac{\partial}{\partial \mathbf{x}} f(\cdot)\right]^T$ denotes the Hessian operator.

II. System Model

We consider a communication system where multiple nodes communicate with a single destination node. Multi-relay cooperative or multi-user distributed networks are examples of the system model under consideration. All nodes are assumed to be equipped with a single antenna. In addition to being affected by multiple channel parameters, the received signal is affected by MTOs due to the random propagation delays at each node.

Signal transmission consists of a *training period* and a *data transmission period*. In the training period, the estimates of MTOs and channel parameters are obtained by employing different TSs of length L that are transmitted from the K distributed nodes. In the data transmission period, the estimates obtained during the training period are used at the receiver to detect the signals from multiple nodes. In order to improve the overall end-to-end system performance in the data transmission period, the goal of this paper is to design optimal TSs in the training period to efficiently estimate MTOs and channel parameters. Throughout this paper, the following set of assumptions and system design parameters are considered:

A1. Quasi-static and frequency-flat fading channels are considered, i.e., the channel parameters do not change over the length of a frame but they change from frame to frame. The assumption of frequency-flat channels can be broadened to frequency-selective channels by employing orthogonal frequency division multiple access. Moreover, the use of such channels is motivated by prior research in this field [4], [25]–[29].

A2. Over a frame, timing offsets are modeled as deterministic but *unknown* parameters [4], [25], [26].

A3. The effect of CFO on the received signal is not considered, since the topic of TS design for MCFO estimation has been extensively addressed in the literature, e.g., see [30] and references therein.

The sampled baseband received signal, $\mathbf{y} \triangleq [y(0),\ldots,y(QL-1)]^T$, at the receiver, prior to matched filtering, is given by

$$\mathbf{y} = \mathbf{\Psi}\mathbf{h} + \mathbf{w}, \tag{1}$$

where:

- T denotes the symbol duration, $T_s = T/Q$ is the sampling interval, Q is the oversampling factor,
- $\mathbf{\Psi} \triangleq [\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_K]$ is a $QL \times K$ matrix, $\boldsymbol{\xi}_k \triangleq \mathbf{G}_k\mathbf{t}_k$ is a $QL \times 1$ vector for $k = 1,\ldots,K$,

- $\mathbf{G}_k \triangleq [\mathbf{g}_0(\tau_k), \ldots, \mathbf{g}_{L-1}(\tau_k)]$ is a $QL \times L$ matrix of the samples of the pulse shaping filter such that $\mathbf{g}_n(\tau_k) \triangleq [g(-nT+\tau_k T), \ldots, g(-nT+iT_s+\tau_k T), \ldots, g(-nT+(QL-1)T_s+\tau_k T)]^T$, for $n = 0, \ldots, L-1$, $\mathbf{t}_k = [t_k(0), \ldots, t_k(L-1)]^T$ denotes the kth node's TS, $\tau_k \in (-0.5, 0.5)$ is the fractional *unknown* timing offset between the kth node and the receiver that is normalized by T,
- $\mathbf{h} \triangleq [h_1, \ldots, h_K]^T$ is the channel vector, h_k denotes the *unknown* channel gain from the kth node to the receiver that changes from frame to frame according to $h_k \sim \mathcal{CN}(0, \sigma_h^2)$, and
- $\mathbf{w} \triangleq [w(0), \ldots, w(QL-1)]^T$ and $w(i)$, for $i = 0, \ldots, QL-1$, denotes the zero-mean complex additive white Gaussian noise (AWGN) at the ith sample of the received signal, i.e., $w(i) \sim \mathcal{CN}(0, \sigma_w^2)$.

Since the root-raise cosine (RRC) waveform is widely applied in communication systems, in this paper, it is used for the pulse shaping filter, $g(t)$ [4], [25], [26]. Furthermore, it is assumed that a coarse synchronizer is first applied and signals from different nodes are within the same symbol period, i.e., the difference between the timing offsets of any two nodes is given by $|\tau_k - \tau_{\bar{k}}| < 1$, for $k, \bar{k} = 1, \ldots, K$ and $k \neq \bar{k}$. Most communication systems use such coarse synchronizers before applying a fine synchronization algorithm to estimate the timing offsets within a symbol period [1], [3], [25]. For clarity, the indices $n, \bar{n} = 0, 1, \ldots, L-1$, $i = 0, 1, \ldots, LQ-1$, and $k, \bar{k} = 1, \ldots, K$ are reserved to denote symbols, T_s-spaced samples, and the nodes, respectively.

III. Estimation Lower Bounds

In this section, closed-form expressions for the HCRB for joint estimation of deterministic MTOs and random channels are derived. Next, to find a tighter lower bound on the variance of estimation error for these parameters, the ECRB is obtained.

A. Hybrid Cramér-Rao Lower Bounds

The HCRB is a lower bound on the joint estimation of random and deterministic parameters and unlike the CRB is not a function of the random parameters [31], [32]. In this context, the derived HCRB does not depend on the particular channels. Thus, by minimizing the HCRB the design guidelines obtained here are general and are applicable to all random realizations of Rayleigh fading channel coefficients. In the following analysis, we make no specific assumption on the distribution of the timing offsets. For example, depending on the receiver design, timing offsets can take on either uniform [25] or normal distributions [22], respectively. To ensure generality, we assume timing offsets to be *deterministic* and unknown parameters that can assume any value within the symbol period. Even though due to this assumption the resulting HCRB is a function of timing offsets, it is numerically shown later that this bound does not vary for different timing offset values. Furthermore, it is also numerically shown in [2] that optimality of TSs is independent of the timing offset values. Hence, the TS design guidelines obtained in the later sections are applicable to the whole range of timing offsets.

The first step in determining the HCRB is to formulate the parameter vector of interest, $\boldsymbol{\theta} \triangleq [\boldsymbol{\theta}_r^T, \boldsymbol{\theta}_d^T]^T$, which is given by

$$\boldsymbol{\theta} \triangleq [\Re\{\mathbf{h}\}^T, \Im\{\mathbf{h}\}^T, \boldsymbol{\tau}^T]^T, \tag{2}$$

where $\boldsymbol{\theta}_r \triangleq [\Re\{\mathbf{h}\}^T, \Im\{\mathbf{h}\}^T]^T$ is the random vector of channel parameters and $\boldsymbol{\theta}_d \triangleq \boldsymbol{\tau} \triangleq [\tau_1, \ldots, \tau_K]^T$ is the deterministic vector of MTOs. Note that according to the assumption in Section II, the complex channel vector is distributed as $\mathbf{h} \sim \mathcal{CN}(\mathbf{0}_{K\times 1}, \sigma_h^2 \mathbf{I}_{K\times K})$. In the following, the *hybrid information matrix* (HIM) and the HCRB for the estimation of $\boldsymbol{\theta}$ are formulated.

Theorem 1: The HIM for estimation of parameters of interest, $\boldsymbol{\theta}$, given the observation vector, $\mathbf{y}$, is a $3K \times 3K$ matrix given by

$$\mathbf{HIM} = \underbrace{\frac{2}{\sigma_w^2}\begin{bmatrix} \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & -\Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \mathbf{0}_{K\times K} \\ \Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \mathbf{0}_{K\times K} \\ \mathbf{0}_{K\times K} & \mathbf{0}_{K\times K} & \Re\{\mathbf{U}\} \end{bmatrix}}_{=\mathbb{E}_{\boldsymbol{\theta}_r|\boldsymbol{\theta}_d}[\mathbf{FIM}]} + \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{\theta}_r}^{-1} & \mathbf{0}_{2K\times K} \\ \mathbf{0}_{K\times 2K} & \mathbf{0}_{K\times K} \end{bmatrix}, \tag{3}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{\theta}_r}^{-1} \triangleq \operatorname{diag}\Big(\underbrace{\frac{2}{\sigma_h^2}, \ldots, \frac{2}{\sigma_h^2}}_{2K}\Big)$ is the $2K \times 2K$ covariance matrix of $\boldsymbol{\theta}_r \triangleq [\Re\{\mathbf{h}\}^T, \Im\{\mathbf{h}\}^T]^T$, $\mathbf{U} \triangleq \sigma_h^2 \operatorname{diag}\big(\boldsymbol{\delta}_1^H\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_K^H\boldsymbol{\delta}_K\big)$ is a $K \times K$ diagonal matrix, $\boldsymbol{\delta}_k \triangleq \mathbf{R}_k\mathbf{t}_k \ \forall\ k$, and $\mathbf{R}_k \triangleq \frac{\partial \mathbf{G}_k}{\partial \tau_k}$ is $QL \times L$ matrix.

Proof: See Appendix A.

HCRB is given by the inverse of the $\mathbf{HIM}$ in (3). In order to ensure that the $\mathbf{HIM}$ in (3) is full rank, the HCRB does not approach infinity, and the parameters of interest can be accurately estimated, it is essential to transmit linearly independent TSs from all the nodes. Moreover, since the off-diagonal blocks of the HIM in (3), i.e., upper right $2K \times K$ and lower left $K \times 2K$ submatrices of $\mathbf{HIM}$, are zero, the HCRB matrix for the estimation of MTOs, $\mathbf{HCRB}(\boldsymbol{\tau})$, is given by

$$\begin{aligned}\mathbf{HCRB}(\boldsymbol{\tau}) &= \frac{\sigma_w^2}{2}\operatorname{diag}\left([\Re\{\mathbf{U}\}]^{-1}\right) \\ &= \frac{\sigma_w^2}{2\sigma_h^2}\left[\frac{1}{\Re\{\boldsymbol{\delta}_1^H\boldsymbol{\delta}_1\}}, \ldots, \frac{1}{\Re\{\boldsymbol{\delta}_K^H\boldsymbol{\delta}_K\}}\right]^T. \end{aligned}\tag{4}$$

Similarly, by using the inverse of the upper left $2K \times 2K$ submatrix of the $\mathbf{HIM}$, the HCRB for estimation of the combined real and imaginary parts of the channel vector is given by (5) at the top of the next page, where $\mathbf{J} = [\mathbf{I}_{K\times K} \quad j\mathbf{I}_{K\times K}]$ is $K \times 2K$ matrix used to obtain the HCRB of $\mathbf{h}$ from the HCRB of $\boldsymbol{\theta}_r$ [31].

B. Conditional Cramér-Rao Lower Bounds

The HCRB derived in Section III-A may not be a tight lower bound on the estimation error variance of an estimator since the Fisher's information matrix, $\mathbf{FIM}$, in (A.3), depends on $\boldsymbol{\theta}$ [31, page 7]. Thus, in this section, the ECRB, which is a tighter lower bound is derived. Following [31, page 6], ECRB is given by

$$\mathbf{ECRB} = \mathbb{E}_{\boldsymbol{\theta}_r}\left[\mathbf{FIM}^{-1}\right], \tag{6}$$

$$\mathbf{HCRB}(\mathbf{h}) = \text{diag}\left(\mathbf{J}\left[\frac{2}{\sigma_w^2}\begin{bmatrix}\Re\{\mathbf{\Psi}^H\mathbf{\Psi}\} & -\Im\{\mathbf{\Psi}^H\mathbf{\Psi}\}\\ \Im\{\mathbf{\Psi}^H\mathbf{\Psi}\} & \Re\{\mathbf{\Psi}^H\mathbf{\Psi}\}\end{bmatrix} + \mathbf{\Sigma}_{\boldsymbol{\theta}_r}^{-1}\right]^{-1}\mathbf{J}^H\right)$$

$$= \frac{\sigma_w^2}{2}\text{diag}\left(\mathbf{J}\underbrace{\begin{bmatrix}\Re\{\mathbf{\Psi}^H\mathbf{\Psi}\} + \frac{\sigma_w^2}{\sigma_h^2}\mathbf{I}_{K\times K} & -\Im\{\mathbf{\Psi}^H\mathbf{\Psi}\}\\ \Im\{\mathbf{\Psi}^H\mathbf{\Psi}\} & \Re\{\mathbf{\Psi}^H\mathbf{\Psi}\} + \frac{\sigma_w^2}{\sigma_h^2}\mathbf{I}_{K\times K}\end{bmatrix}^{-1}}_{\triangleq\mathbf{\Gamma}}\mathbf{J}^H\right). \quad (5)$$

In (6), the expectation is taken with respect to the priori distribution of $\boldsymbol{\theta}_r$, $p(\boldsymbol{\theta}_r) = \frac{\exp\{-\mathbf{h}^H\mathbf{\Sigma}_\mathbf{h}^{-1}\mathbf{h}\}}{(\pi)^K\det(\mathbf{\Sigma}_\mathbf{h})}$ and $\mathbf{\Sigma}_\mathbf{h} \triangleq \text{diag}(\underbrace{\sigma_h^2,\ldots,\sigma_h^2}_{K})$ is the covariance matrix of $\mathbf{h}$. Using the closed-form results for the inverse of the Fisher's information matrix, $\mathbf{FIM}^{-1}$, given in [25], the ECRB for joint estimation of MTOs and multiple channel gains can be determined as (7), given at the bottom of this page, where $\mathbf{D} \triangleq \text{diag}(h_1,\ldots,h_K)$ is a $K\times K$ diagonal matrix and $\mathbf{\Delta} \triangleq \frac{\partial\mathbf{\Psi}}{\partial\boldsymbol{\tau}} \triangleq [\boldsymbol{\delta}_1,\ldots,\boldsymbol{\delta}_K]$ is a $QL\times K$ matrix. It is not mathematically tractable to find closed-form expressions for $\mathbf{ECRB}(\boldsymbol{\tau})$ and $\mathbf{ECRB}(\mathbf{h})$ in (7). Therefore, in this paper, the expectation with respect to $\boldsymbol{\theta}$ in (7) is numerically calculated over a large number of simulated realizations. Using Jensen's inequality [31] and the fact that the matrix, $\mathbf{FIM}$, in (A.3), depends on $\boldsymbol{\theta}$, the relation between ECRB and HCRB can be determined as [31, page 7]

$$\mathbf{ECRB}(\boldsymbol{\tau}) > \mathbf{HCRB}(\boldsymbol{\tau}) \quad \text{and}$$
$$\mathbf{ECRB}(\mathbf{h}) > \mathbf{HCRB}(\mathbf{h}). \quad (8)$$

The above relationships can also be numerically observed through simulations in Fig. 1, where HCRB and ECRB for MTOs estimation are evaluated using (4) and (7a), respectively, for different signal-to-noise-ratios (SNRs). The results in Fig. 1 are obtained for $K = 4$ nodes, TS length $L = 64$, oversampling factor $Q = 2$, and random TSs, i.e., $\mathbf{t}_k = [\exp(-j\phi_0),\ldots,\exp(-j\phi_{L-1})]^T$, $\forall\ k$, and $\phi_n \sim \mathcal{U}(-\pi,\pi)$, $\forall\ n$. Without loss of generality, only the lower bounds for the first node are plotted for different values of timing offsets, $\tau = [-0.5 - 0.4,\ldots,0.5]$. Moreover, the HCRB and the ECRB values are averaged over 500 trials, where for each trial, the timing offsets for the remaining nodes are assumed to be uniformly distributed over $(-0.5, 0.5)$. It can be observed from Fig. 1 that the HCRB is a lower bound compared to the ECRB for both SNR = 10 dB and SNR = 20 dB. This outcome is also confirmed for different TSs in Section VI. Fig. 1 also shows that in the presence of oversampling, both the HCRB and the ECRB do not vary for different timing offset values, which is also confirmed later for all the proposed TSs in Section VI. Consequently, although the HCRB is a function of timing offsets, the TS design guidelines obtained by minimizing the HCRB in Section IV are independent of the timing offset values.

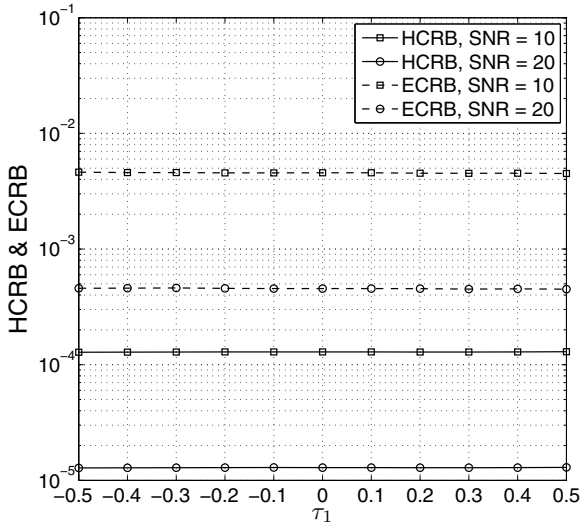


Fig. 1. HCRB and ECRB for the estimation of τ_1 for different values of τ_1 with $K = 4$ nodes.

Recall that even though the ECRB is a tighter bound compared to the HCRB, it cannot be derived in closed-form. Thus, analytical solutions for the optimal TSs cannot be obtained by minimizing the ECRB. As a result, in Section IV, guidelines for optimal TS design are obtained by minimizing the HCRB instead. Nevertheless, in Section VI, through numerical simulations it is demonstrated that the TSs that minimize the HCRB also minimize the ECRB and the MSE of the derived MAP estimator.

IV. Training Sequence Design

In this paper, the optimal TS is defined as the TS that jointly minimizes the HCRBs of $\boldsymbol{\tau}$ and $\mathbf{h}$. The following subsections present criteria required for minimizing the HCRBs of $\boldsymbol{\tau}$ and $\mathbf{h}$.

$$\mathbf{ECRB}(\boldsymbol{\tau}) = \frac{\sigma_w^2}{2}\mathbb{E}_{\boldsymbol{\theta}_r}\left[\left(\Re\left\{\mathbf{D}^H\mathbf{\Delta}^H\left(\mathbf{I}_{LQ\times LQ} - \mathbf{\Psi}(\mathbf{\Psi}^H\mathbf{\Psi})^{-1}\mathbf{\Psi}^H\right)\mathbf{\Delta}\mathbf{D}\right\}\right)^{-1}\right], \quad (7a)$$

$$\mathbf{ECRB}(\mathbf{h}) = \frac{\sigma_w^2}{2}\mathbb{E}_{\boldsymbol{\theta}_r}\Big[2(\mathbf{\Psi}^H\mathbf{\Psi})^{-1} + (\mathbf{\Psi}^H\mathbf{\Psi})^{-1}\mathbf{\Psi}^H\mathbf{\Delta}\mathbf{D}\left(\Re\left\{\mathbf{D}^H\mathbf{\Delta}^H\left(\mathbf{I}_{LQ\times LQ} - \mathbf{\Psi}(\mathbf{\Psi}^H\mathbf{\Psi})^{-1}\mathbf{\Psi}^H\right)\mathbf{\Delta}\mathbf{D}\right\}\right)^{-1}$$
$$\mathbf{D}^H\mathbf{\Delta}^H\mathbf{\Psi}(\mathbf{\Psi}^H\mathbf{\Psi})^{-1}\Big]. \quad (7b)$$

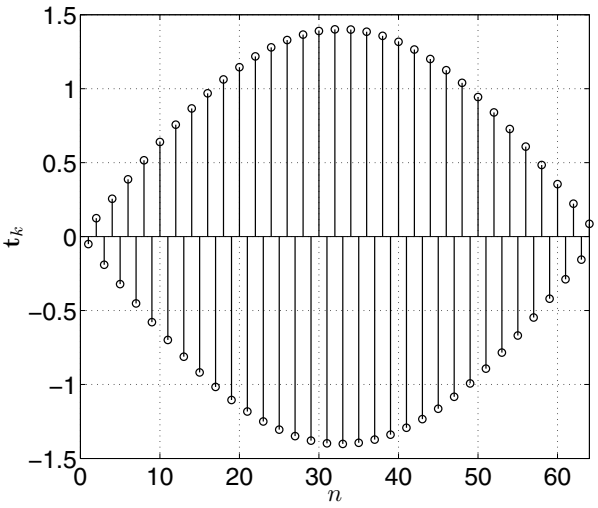


Fig. 2. Optimal solution to (9) for the minimization of $\mathbf{HCRB}(\boldsymbol{\tau})$, where TS length $L = 64$ and $\mathbf{R}_k$ is evaluated at $\tau_k = 0.4$.

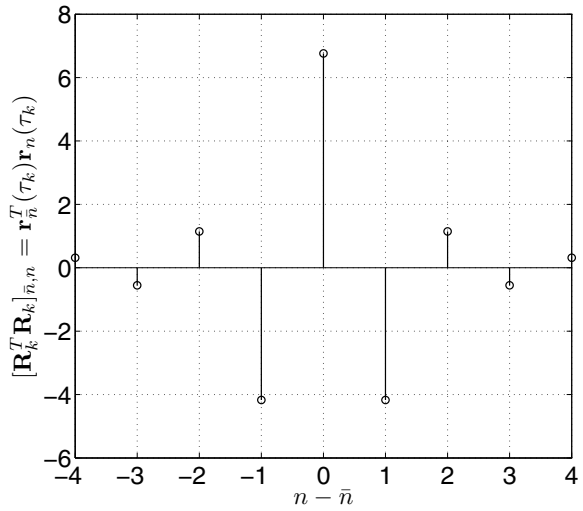


Fig. 3. $[\mathbf{R}_k^T\mathbf{R}_k]_{\bar{n},n} = \mathbf{r}_{\bar{n}}^T(\tau_k)\mathbf{r}_n(\tau_k)$ for different values of $n - \bar{n}$ with $\tau_k = 0.4$, $n = L/2$ and $L = 64$.

A. *Minimization of* $\mathbf{HCRB}(\boldsymbol{\tau})$

To minimize $\mathbf{HCRB}(\boldsymbol{\tau})$ in (4), we have to maximize each of $\Re\{\boldsymbol{\delta}_1^H\boldsymbol{\delta}_1\}, \ldots, \Re\{\boldsymbol{\delta}_K^H\boldsymbol{\delta}_K\}$. Using the definition of $\boldsymbol{\delta}_k \triangleq \mathbf{R}_k\mathbf{t}_k$ below (3), for $k = 1, \ldots, K$, the optimal TS, $\mathbf{t}_k$, that minimizes the HCRB of $\boldsymbol{\tau}$, is the solution to

$$\arg\max_{\mathbf{t}_k} \quad \Re\{\mathbf{t}_k^H\mathbf{R}_k^T\mathbf{R}_k\mathbf{t}_k\}, \qquad \mathtt{s.t.} \quad \mathbf{t}_\mathtt{k}^\mathtt{H}\mathbf{t}_\mathtt{k} \leq \mathtt{L}, \tag{9}$$

where $\mathbf{R}_k \triangleq \frac{\partial \mathbf{G}_k}{\partial \tau_k}$ is a matrix of real numbers, since elements of matrix $\mathbf{G}_k$ are samples of the real RRC waveform. Given that $\mathbf{x}^H\mathbf{R}_k^T\mathbf{R}_k\mathbf{x} = \|\mathbf{R}_k\mathbf{x}\|^2 \geq 0$ for any $L \times 1$ vector $\mathbf{x}$ and all eigenvalues of $\mathbf{R}_k^T\mathbf{R}_k$ are greater than zero, $\mathbf{R}_k^T\mathbf{R}_k$ is a symmetric positive definite matrix and the optimization problem in (9) is convex [33]. Thus, the optimal solution to (9) is given by

$$\mathbf{t}_k = \sqrt{L}\lambda_{\max}\left(\mathbf{R}_k^T\mathbf{R}_k\right), \tag{10}$$

where $\lambda_{\max}\left(\mathbf{R}_k^T\mathbf{R}_k\right)$ is the eigenvector corresponding to the maximum eigenvalue of $\mathbf{R}_k^T\mathbf{R}_k$. The resulting TS is shown in Fig. 2 where $\mathbf{R}_k$ is evaluated at $\tau_k = 0.4$ and TS length $L = 64$. It is important to note that the TS in Fig. 2 undergoes a sign change from symbol to symbol and similar TSs, as shown in Fig. 2, are also obtained by evaluating (10) for different values of τ_k.

Remark 1: To develop more insight into the optimal solution in (10) and establish a more comprehensive TS design guideline that maximizes the cost function in (9), we numerically study the structure of the matrix $\mathbf{R}_k$. Let us write $\mathbf{R}_k$ in terms of its column vectors, $\mathbf{R}_k = [\mathbf{r}_0(\tau_k), \ldots, \mathbf{r}_{L-1}(\tau_k)]$, where $\mathbf{r}_n(\tau_k) = [r(-nT + \tau_k T), \ldots, r(-nT + iT_s + \tau_k T), \ldots, r(-nT + (QL-1)T_s + \tau_k T)]^T$, $\forall\ n$, and $r(t) \triangleq \frac{\partial g(t)}{\partial \tau_k}$. $\mathbf{R}_k^T\mathbf{R}_k$ is an $L \times L$ matrix such that

$$\left[\mathbf{R}_k^T\mathbf{R}_k\right]_{\bar{n},n} = \mathbf{r}_{\bar{n}}^T(\tau_k)\mathbf{r}_n(\tau_k). \tag{11}$$

Fig. 3 plots the elements of matrix $\mathbf{R}_k^T\mathbf{R}_k$, e.g., diagonal elements $(n - \bar{n} = 0)$, the first diagonal above $(n - \bar{n} = 1)$ and below the main diagonal $(n - \bar{n} = -1)$, the second diagonal above $(n - \bar{n} = 2)$ and below the main diagonal $(n - \bar{n} = -2)$, and so on for $n = L/2$, $L = 64$, and $\tau_k = 0.4$. It can be observed from Fig. 3 that for the matrix $\mathbf{R}_k^T\mathbf{R}_k$, the elements corresponding to even values of $(n - \bar{n})$ are positive while the elements for the odd values of $(n - \bar{n})$ are negative[1]. Similar results are also observed $\forall\ n, \tau_k$. Consequently, the cost function in (9), $\Re\{\mathbf{t}_k^H\mathbf{R}_k^T\mathbf{R}_k\mathbf{t}_k\}$, is maximized, when the transmitted TS alternates in sign every symbol period since based on the structure of $\mathbf{R}_k^T\mathbf{R}_k$, opposite-signed TS symbols are multiplied by the negative-valued elements of $\mathbf{R}_k^T\mathbf{R}_k$. Similarly, it can be concluded that for complex TSs, the cost function, $\Re\{\mathbf{t}_k^H\mathbf{R}_k^T\mathbf{R}_k\mathbf{t}_k\}$ is maximized by transmitting the TS that exhibits a phase shift of π radians from symbol to symbol. Fig. 3 also shows that the magnitude of elements $\mathbf{R}_k^T\mathbf{R}_k$ decay as $|n - \bar{n}|$ increases, since the RRC function decays with every sample.

The above analysis shows that TSs that undergo a π radian phase shift every symbol period are consistent with the optimal solution that minimizes the HCRB for MTO estimation. Such TSs with sign changes are also reported as optimal based on intuition or simulation in [2], [24]. However, no analytical results were included to support these claims in [2] and [24].

In [2], it is shown that the TS in Fig. 2 is optimal for estimation of a *single* timing offset in point-to-point MIMO systems. In this paper, our focus is to design optimal TSs for joint estimation of MTO and multiple channel parameters, which will be obtained by jointly minimizing $\mathbf{HCRB}(\boldsymbol{\tau})$ and $\mathbf{HCRB}(\mathbf{h})$.

B. *Minimization of* $\mathbf{HCRB}(\mathbf{h})$

To minimize $\mathbf{HCRB}(\mathbf{h})$ in (5), a closed-form expression for the matrix inverse, $\boldsymbol{\Gamma}$, in (5) needs to be determined. For asymptotically large values of TS length, L,

$$\mathrm{diag}(\Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\}) \gg \mathrm{diag}\left(\frac{\sigma_w^2}{\sigma_h^2}\mathbf{I}_{K\times K}\right), \tag{12}$$

[1]This is also because the slope of the RRC function, $r(t)$, changes after every time period T.

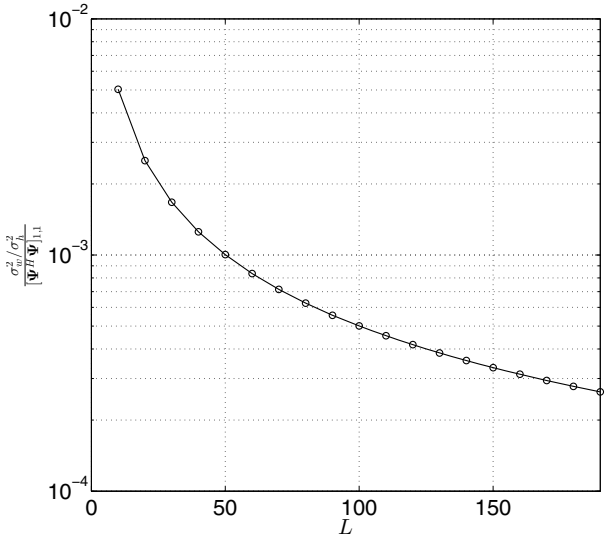


Fig. 4. $\frac{\sigma_w^2/\sigma_h^2}{[\mathbf{\Psi}^H\mathbf{\Psi}]_{1,1}}$ versus TS length, L, for oversampling factor $Q = 2$, SNR = 10 dB ($\sigma_w^2 = 0.1$), and $\sigma_h^2 = 1$.

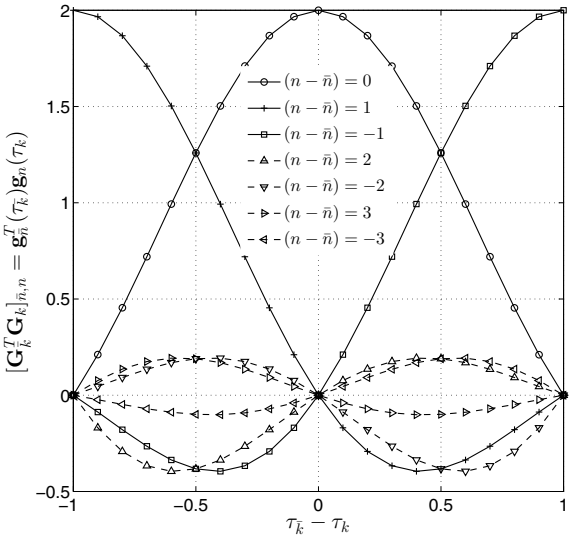


Fig. 5. $[\mathbf{G}_{\bar{k}}^T\mathbf{G}_k]_{\bar{n},n} = \mathbf{g}_{\bar{n}}^T(\tau_{\bar{k}})\mathbf{g}_n(\tau_k)$ for different values of τ_k and $\tau_{\bar{k}}$, such that $|\tau_{\bar{k}} - \tau_k| < 1$, $n = L/2$ and $L = 64$.

since for large values of L the contribution of the prior information matrix to the HCRB is considerably smaller than $\mathbb{E}_{\boldsymbol{\theta}}[\mathbf{FIM}]$ [31]. This is also numerically shown in Fig. 4, where $\frac{\sigma_w^2/\sigma_h^2}{[\mathbf{\Psi}^H\mathbf{\Psi}]_{1,1}}$ is plotted versus TS length, L, for oversampling factor $Q = 2$, SNR = 10 dB ($\sigma_w^2 = 0.1$), and $\sigma_h^2 = 1$. Random TSs are used similar to those of Section III-B.[2]

Using (12), (5) can be approximated as

$$\widetilde{\mathbf{HCRB}}(\mathbf{h}) = \frac{\sigma_w^2}{2}\mathrm{diag}\left(\mathbf{J}\begin{bmatrix} \Re\{\mathbf{\Psi}^H\mathbf{\Psi}\} & -\Im\{\mathbf{\Psi}^H\mathbf{\Psi}\} \\ \Im\{\mathbf{\Psi}^H\mathbf{\Psi}\} & \Re\{\mathbf{\Psi}^H\mathbf{\Psi}\}\end{bmatrix}^{-1}\mathbf{J}^H\right) = \sigma_w^2\mathrm{diag}\Big([\mathbf{\Psi}^H\mathbf{\Psi}]^{-1}\Big), \quad (13)$$

where $\widetilde{\mathbf{HCRB}}(\mathbf{h})$ is the asymptotic HCRB for estimation of channel, $\mathbf{h}$. According to (13), minimizing $\widetilde{\mathbf{HCRB}}(\mathbf{h})$ is equivalent to minimizing $\mathrm{Tr}\left([\mathbf{\Psi}^H\mathbf{\Psi}]^{-1}\right)$. Based on the results in [34, page 65], the following lemma applies.

Lemma 1: For a $K \times K$ positive definite matrix $\mathbf{X}$, $\mathrm{Tr}\left(\mathbf{X}^{-1}\right) \geq \sum_{k=1}^{K}\frac{1}{[\mathbf{X}]_{k,k}}$, with equality sign applying if $\mathbf{X}$ is diagonal.

Using Lemma 1, for the positive definite matrix $\mathbf{\Psi}^H\mathbf{\Psi}$, we have

$$\mathrm{Tr}\left([\mathbf{\Psi}^H\mathbf{\Psi}]^{-1}\right) \geq \sum_{k=1}^{K}\frac{1}{[\mathbf{\Psi}^H\mathbf{\Psi}]_{k,k}}. \quad (14)$$

It is shown in Appendix B that $\mathbf{\Psi}^H\mathbf{\Psi}$ is a positive definite matrix. Thus, using (14), HCRB of $\mathbf{h}$ can be minimized by ensuring that the off diagonal elements of the $K \times K$ matrix $\mathbf{\Psi}^H\mathbf{\Psi}$, whose elements are given by

$$[\mathbf{\Psi}^H\mathbf{\Psi}]_{\bar{k},k} = \mathbf{t}_{\bar{k}}^H\mathbf{G}_{\bar{k}}^T\mathbf{G}_k\mathbf{t}_k, \qquad k,\bar{k} = 1,\ldots,K, \quad (15)$$

are zero, i.e., $\mathbf{\Psi}^H\mathbf{\Psi}$ is a diagonal matrix. In order to minimize the off-diagonal elements of $\mathbf{\Psi}^H\mathbf{\Psi}$, the structure of the $L \times L$ matrix $\mathbf{G}_{\bar{k}}^T\mathbf{G}_k$, which depends on the values of timing offsets from different nodes needs to be analyzed. Let us consider the diagonal elements of $\mathbf{\Psi}^H\mathbf{\Psi}$ first. Since after a symbol period T, the RRC function decays very quickly, it is observed numerically that the matrix $\mathbf{G}_k^T\mathbf{G}_k$ is diagonally dominant[3] for any value of τ_k. Consequently, the diagonal elements of $\mathbf{\Psi}^H\mathbf{\Psi}$ that are given by (15), with $k = \bar{k}$, are maximized if $\mathbf{t}_k^H\mathbf{t}_k = L$.

The off-diagonal elements of $\mathbf{\Psi}^H\mathbf{\Psi}$ are given by (15), for $k \neq \bar{k}$, which depend on the structure of the matrix $\mathbf{G}_{\bar{k}}^T\mathbf{G}_k$, for $k \neq \bar{k}$. It is numerically observed that $\mathbf{G}_{\bar{k}}^T\mathbf{G}_k$ is a tri-diagonally dominant matrix for any value of τ_k and $\tau_{\bar{k}}$ with $|\tau_{\bar{k}} - \tau_k| < 1$. This is also numerically demonstrated in Fig. 5, where the elements of $\mathbf{G}_{\bar{k}}^T\mathbf{G}_k$, $[\mathbf{G}_{\bar{k}}^T\mathbf{G}_k]_{n,\bar{n}} = \mathbf{g}_n(\tau_{\bar{k}})^T\mathbf{g}_{\bar{n}}(\tau_k)$, are plotted for different values of $|\tau_{\bar{k}} - \tau_k|$ and $n - \bar{n}$. Fig. 5 shows that given $|\tau_{\bar{k}} - \tau_k| < 1$, the diagonal elements, $n - \bar{n} = 0$, the first diagonal elements above the main diagonal, $n - \bar{n} = 1$, and the first diagonal elements below the main diagonal, $n - \bar{n} = -1$, of matrix $\mathbf{G}_{\bar{k}}^T\mathbf{G}_k$ are dominant. Consequently, to minimize the off-diagonal elements of $\mathbf{\Psi}^H\mathbf{\Psi}$, i.e., $[\mathbf{\Psi}^H\mathbf{\Psi}]_{\bar{k},k} = \mathbf{t}_{\bar{k}}^H\mathbf{G}_{\bar{k}}^T\mathbf{G}_k\mathbf{t}_k$, for $k \neq \bar{k}$, the following two conditions can be developed:

1) The TSs from different nodes need to be mutually orthogonal.
2) The TSs from any node should also be mutually orthogonal to $+T$-shifted TSs and $-T$-shifted TSs from every other node to minimize effects on the first diagonals above and below the main diagonal elements of $\mathbf{G}_{\bar{k}}^T\mathbf{G}_k$, respectively.

[2]Though results for only first node are presented, similar results are observed for the remaining nodes.

[3]These results are not presented here due to lack of space.

TABLE I
TSs IN PROPOSED EXAMPLE 1 FOR $K = 4$ RELAYS AND TS LENGTH $L = 64$.

$\mathbf{t}_1$	$[-0.05, 0.12, -0.19, 0.25, -0.32, 0.38, -0.45, 0.51, -0.57, 0.63, -0.69, 0.75, -0.81, 0.86, -0.91, 0.96, -1.01, 1.06, -1.10, 1.14, -1.18, 1.21, -1.24, 1.27, -1.30, 1.32, -1.34, 1.36, -1.37, 1.38, -1.39, 1.40, -1.40, 1.39, -1.39, 1.38, -1.37, 1.35, -1.33, 1.31, -1.29, 1.26, -1.23, 1.20, -1.16, 1.12, -1.08, 1.03, -0.99, 0.94, -0.89, 0.83, -0.78, 0.72, -0.66, 0.60, -0.54, 0.48, -0.41, 0.35, -0.28, 0.22, -0.15, 0.08]^T$
$\mathbf{t}_2$	$[0.10, -0.24, 0.37, -0.50, 0.62, -0.74, 0.85, -0.96, 1.05, -1.13, 1.21, -1.27, 1.32, -1.36, 1.38, -1.40, 1.40, -1.38, 1.36, -1.32, 1.27, -1.20, 1.13, -1.04, 0.95, -0.84, 0.73, -0.62, 0.49, -0.36, 0.23, -0.10, -0.03, 0.16, -0.30, 0.43, -0.55, 0.68, -0.79, 0.90, -1.00, 1.09, -1.17, 1.24, -1.29, 1.34, -1.37, 1.39, -1.40, 1.39, -1.37, 1.34, -1.29, 1.24, -1.17, 1.09, -1.00, 0.90, -0.80, 0.68, -0.56, 0.43, -0.30, 0.17]^T$
$\mathbf{t}_3$	$[-0.15, 0.37, -0.55, 0.73, -0.89, 1.04, -1.16, 1.26, -1.34, 1.38, -1.40, 1.38, -1.34, 1.27, -1.17, 1.05, -0.91, 0.74, -0.57, 0.37, -0.18, -0.02, 0.22, -0.42, 0.60, -0.78, 0.94, -1.08, 1.20, -1.29, 1.35, -1.39, 1.40, -1.37, 1.32, -1.24, 1.14, -1.01, 0.86, -0.69, 0.51, -0.32, 0.12, 0.08, -0.28, 0.47, -0.66, 0.83, -0.98, 1.12, -1.23, 1.31, -1.37, 1.39, -1.39, 1.36, -1.30, 1.22, -1.10, 0.97, -0.81, 0.64, -0.45, 0.25]^T$
$\mathbf{t}_4$	$[-0.20, 0.48, -0.72, 0.93, -1.11, 1.25, -1.35, 1.40, -1.39, 1.33, -1.22, 1.06, -0.86, 0.64, -0.39, 0.12, 0.14, -0.40, 0.65, -0.88, 1.07, -1.22, 1.33, -1.39, 1.39, -1.35, 1.25, -1.11, 0.92, -0.70, 0.46, -0.20, -0.06, 0.33, -0.59, 0.82, -1.02, 1.19, -1.30, 1.38, -1.40, 1.36, -1.28, 1.15, -0.98, 0.77, -0.53, 0.27, -0.005, -0.26, 0.52, -0.76, 0.97, -1.14, 1.28, -1.36, 1.4, -1.38, 1.31, -1.19, 1.03, -0.83, 0.60, -0.33]^T$

C. Conditions for Optimal TSs and Proposed Training Sequences

In summary, from Section IV-A and Section IV-B, the following three conditions can be obtained for the design of optimal TSs:

C1: The TSs from all nodes exhibit a phase shift of π radians every symbol period, which corresponds to (10).
C2: The TSs from all nodes be mutually orthogonal.
C3: The TSs from any node be orthogonal to $\pm\, T$-shifted TSs from every other node.

As highlighted in Section I, the design guidelines in [24] are not complete. In this paper, according to the proposed analytical framework, a new criterion, C3, is proposed. The results in Section VI demonstrate that neglecting C3 in the design of TSs significantly degrades joint channel and MTO estimation performance. Moreover, guideline C1 is generalized for complex TSs. It is not easy to find the TSs that jointly satisfy C1-C3, since the TSs that satisfy C2 and C3 do not exactly satisfy C1 and vice versa. However, in what follows, we propose two techniques for generating TSs that nearly satisfy the above conditions and improve estimation and relaying performance.

1) Proposed Technique 1 (Proposed-1): Although the TSs obtained by optimizing the cost function in (9) exhibit a π-radian phase shift every symbol period, they do not meet conditions C2 and C3. To satisfy C2, we propose to exploit the K eigenvectors of $\mathbf{R}_k^T\mathbf{R}_k$ since they are mutually orthogonal. Consequently, K TSs can be obtained by multiplying the K eigenvectors corresponding to the first K maximum eigenvalues of $\left(\mathbf{R}^{[0]}\right)^T\mathbf{R}^{[0]}$, where $\mathbf{R}^{[0]} \triangleq \left.\frac{\partial \mathbf{G}}{\partial \tau}\right|_{\tau=0}$, by $\sqrt{L}$. This technique can be used to construct TSs of any desired length. For example, the TSs for $K = 4$ nodes and TS length $L = 64$ are shown in Table I. It can be observed from Table I that all 4 TSs exhibit π radian phase shift every symbol period satisfying C1. In addition, C2 is also satisfied because eigenvectors of $\left(\mathbf{R}^{[0]}\right)^T\mathbf{R}^{[0]}$ are mutually orthogonal. Moreover, even though C3 is not exactly satisfied, every TS is nearly orthogonal to $\pm T$ shifted TSs from every other node and the numerical results in Section VI show that the proposed TSs significantly enhance estimation and system performance. Note that the Proposed-1 TS matches with the TS proposed in [2] for single timing offset estimation in MIMO systems.

2) Proposed Technique 2 (Proposed-2): We propose to use the Walsh-Hadamard codes as a second approach for generating optimal TSs since they always meet condition C2 [35]. The Walsh-Hadamard codes here are generated using the initial matrix $\mathbf{W} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, which is repeated $\log_2(L)-1$ times according to $\mathbf{W} = \begin{bmatrix} \mathbf{W} & \mathbf{W} \\ \mathbf{W} & -\mathbf{W} \end{bmatrix}$ to get an $L \times L$ Hadamard matrix, $\mathbf{W}$, where the TS length in this case is $L = 2^q$, for any positive integer q. To satisfy C1, we propose to select the columns of $\mathbf{W}$ as TSs that exhibit maximum phase shift from symbol to symbol, i.e., $L-1, \ldots, L-K$ sign changes for K nodes. Similar to the technique above, the proposed TSs do not fully satisfy conditions C1 and C3 but they are nearly optimal and the results in Section VI show that they considerably enhance system performance.

In (9), the necessary and sufficient condition for the design of training sequences (TSs) that minimize $\mathbf{HCRB}(\boldsymbol{\tau})$ is provided. In (10) the solution for meeting this condition is also derived. Moreover, in (13) and Lemma 1, for asymptotically large TS length, L, the necessary and sufficient condition for minimizing $\mathbf{HCRB}(\mathbf{h})$ is also established. Although based on these conditions the TS design guidelines C1–C3 are derived, it cannot be analytically established that C1–C3 are also necessary conditions to minimize $\mathbf{HCRB}(\boldsymbol{\tau})$ and $\mathbf{HCRB}(\mathbf{h})$. Nevertheless, our simulation results in Section VI indicate that the proposed conditions C1–C3 are necessary conditions for optimal TS design because the proposed TSs based on C1–C3 achieve a BER performance that is close to the benchmark BER. Moreover, the TSs that violate these conditions demonstrate poor BER performance.

V. MAXIMUM-A-POSTERIORI ESTIMATOR

In this section, a *data-aided* MAP estimator for joint estimation of MTOs and multiple channels is derived. The MAP estimate of $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}$ maximizes the log of the posterior probability density function (PDF), given by

$$\log p(\boldsymbol{\theta}|\mathbf{y}) = \log\left\{\frac{p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{y})}\right\}, \tag{16}$$

where $p(\mathbf{y}|\boldsymbol{\theta})$ is the likelihood function given by

$$p(\mathbf{y}|\boldsymbol{\theta}) = \frac{1}{(\pi\sigma_w^2)^{LQ}}\exp\left\{-\frac{\|\mathbf{y}-\boldsymbol{\Psi}\mathbf{h}\|^2}{\sigma_w^2}\right\}, \tag{17}$$

TABLE II
SUMMARY OF TSS USED IN THE SIMULATIONS WITH TS LENGTH $L = 64$ AND $K = 4$ NODES.

Proposed-1	$\sqrt{L}$ times the eigenvectors corresponding to the first 4 maximum eigenvalues of $\left(\mathbf{R}^{[0]}\right)^T \mathbf{R}^{[0]}$ $\left(\mathbf{R}^{[0]}\right.$ is defined in Sec. IV-C1 $\left.\right)$
Proposed-2	Walsh-Hadamard Matrix $\mathbf{W}$ given in Sec. IV-C2 (Column No. 2, 18, 34, and 50)
Violate-C1	Walsh-Hadamard Matrix $\mathbf{W}$ given in Sec. IV-C2 (Column No. 1, 17, 33, and 49)
Violate-C2	$\mathbf{t}_1 = [1.00 + j0.00, -0.92 + j0.38, 0.70 - j0.70, -0.38 + j0.92, \ldots]^T$ $\mathbf{t}_2 = [0.92 + 0.38, -0.70 - j0.70, 0.38 + j0.92, 0.00 - j1.00, \ldots]^T$ $\mathbf{t}_3 = [0.70 + j0.70, 0.00 - 1.00, 0.70 - j0.70, 0.00 + j1.00, \ldots]^T$ $\mathbf{t}_4 = [0.00 + j1.00, 0.00 - j1.00, 0.00 + j1.00, 0.00 - j1.00, \ldots]^T$
Violate-C3	Walsh-Hadamard Matrix $\mathbf{W}$ given in Sec. IV-C2 (Column No. 5, 6, 7, and 8)

and $p(\boldsymbol{\theta})$ is the priori distribution of $\boldsymbol{\theta}$, given by

$$p(\boldsymbol{\theta}) = \frac{1}{(\pi)^K \det(\boldsymbol{\Sigma}_{\mathbf{h}})} \exp\left\{-\mathbf{h}^H \boldsymbol{\Sigma}_{\mathbf{h}}^{-1} \mathbf{h}\right\}. \tag{18}$$

In (18), $\boldsymbol{\Sigma}_{\mathbf{h}} \triangleq \text{diag}\big(\underbrace{\sigma_h^2, \ldots, \sigma_h^2}_{K}\big)$ is the covariance matrix of $\mathbf{h}$ and $\boldsymbol{\Sigma}_{\mathbf{h}}^{-1} \triangleq \text{diag}\Big(\underbrace{\frac{1}{\sigma_h^2}, \ldots, \frac{1}{\sigma_h^2}}_{K}\Big)$. Substituting (17) and (18) into (16), the posterior PDF in (16) can be written as

$$\log p(\boldsymbol{\theta}|\mathbf{y}) = -\log(\pi\sigma_h^2)^K (\pi\sigma_w^2)^{LQ} - \frac{\|\mathbf{y} - \boldsymbol{\Psi}\mathbf{h}\|^2}{\sigma_w^2} - \mathbf{h}^H \boldsymbol{\Sigma}_{\mathbf{h}}^{-1} \mathbf{h} - \log p(\mathbf{y}), \tag{19}$$

where $\log(\pi\sigma_h^2)^K (\pi\sigma_w^2)^{LQ}$ and $\log p(\mathbf{y})$ are independent of $\boldsymbol{\theta}$. Thus, the MAP estimate of $\boldsymbol{\theta}$ is given by

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \left\{ \frac{\|\mathbf{y} - \boldsymbol{\Psi}\mathbf{h}\|^2}{\sigma_w^2} + \mathbf{h}^H \boldsymbol{\Sigma}_{\mathbf{h}}^{-1} \mathbf{h} \right\}. \tag{20}$$

Taking the derivative of the above cost function with respect to $\mathbf{h}^H$ and equating the result to zero, the MAP estimate of channels, $\hat{\mathbf{h}}$, is determined as

$$\hat{\mathbf{h}} = \left(\boldsymbol{\Psi}^H \boldsymbol{\Psi} + \boldsymbol{\Sigma}_{\mathbf{h}}^{-1}\right)^{-1} \boldsymbol{\Psi}^H \mathbf{y}. \tag{21}$$

Substituting (21) back into the cost function in (20), the MAP estimate of MTOs, $\hat{\boldsymbol{\tau}}$, is determined as

$$\hat{\boldsymbol{\tau}} = \arg\max_{\boldsymbol{\tau}} \left\{ \mathbf{y}^H \boldsymbol{\Psi} \left(\boldsymbol{\Psi}^H \boldsymbol{\Psi} + \boldsymbol{\Sigma}_{\mathbf{h}}^{-1}\right)^{-1} \boldsymbol{\Psi}^H \mathbf{y} \right\}. \tag{22}$$

Using the estimate of $\hat{\boldsymbol{\tau}}$ from (22) and by evaluating $\boldsymbol{\Psi}$ at $\boldsymbol{\tau} = \hat{\boldsymbol{\tau}}$, the channel estimates, $\hat{\mathbf{h}}$, can be obtained using (21). Since $\mathbf{y}$ and $\boldsymbol{\theta}$ are not jointly Gaussian, it cannot be analytically shown that the proposed MAP estimator in (21) and (22) is unbiased [34]. However, through simulations, we have found that the estimated parameters in (21) and (22) have a very small bias, e.g., at SNR $= 10$ dB, it is found that $\mathbb{E}\{\boldsymbol{\tau} - \hat{\boldsymbol{\tau}}\} = -0.0027$ and $\mathbb{E}\{\mathbf{h} - \hat{\mathbf{h}}\} = 0.0009 - j0.0013$, when averaged over 500 simulations. In the following section, the estimation accuracy of the proposed MAP estimator will be assessed by the ECRBs, $\mathbf{ECRB}(\boldsymbol{\tau})$ and $\mathbf{ECRB}(\mathbf{h})$, given in (7).

VI. SIMULATION RESULTS

In order to demonstrate the advantage of the proposed TSs for improving estimation and system performance, they are compared to TSs that violate the proposed design guidelines, C1-C3. The two proposed TSs and three non-optimal TSs, which violate C1, C2, and C3, respectively, are given in Table II for TS length $L = 64$. Throughout this section, the proposed TSs in Sections IV-C1 and IV-C2 are referred to as Proposed-1 and Proposed-2. The three non-optimal TSs are referred to as Violate-C1, Violate-C2, and Violate-C3 and are chosen as follows.

Walsh-Hadamard codes are used for the non-optimal TSs denoted by Violate-C1. Even though these TSs satisfy C2, they are selected in a fashion to not satisfy C1, i.e., have the least frequent sign changes. As shown in Section IV-A this choice of TSs is expected to increase the HCRB for estimation of $\boldsymbol{\tau}$. The next TSs referred to as Violate-C2, are generated using 16-phase-shift keying (16-PSK) modulated symbols. It is not possible to completely satisfy C1 by having symbol-to-symbol phase shifts of π radians for all the TSs since the resulting TSs will be linearly dependent and the $\mathbf{HIM}$ in (3) will be ill-conditioned. However, in order to comply with C1 as closely as possible, the TSs are selected to have maximum phase shift, close to π radians, e.g., π, $7\pi/8$, $9\pi/8$, $6\pi/8$ radians, from symbol to symbol. Moreover, the resulting TSs are correlated and violate C2 to a large extent, which according to the results in Section IV-B, increases the HCRB for channel estimation. Finally, for Violate-C3 TSs, the Walsh-Hadamard codes are applied again, satisfying C2. However, these codes are selected to ensure that C3 is mainly violated, i.e., the inner product of $\mathbf{t}_k$ and ± 1 shifted $\mathbf{t}_{\bar{k}}$, for $k \neq \bar{k}$, is large. As shown in Section IV-B, this also results in a larger HCRB for estimation of channels, $\mathbf{h}$.

Without loss of generality, sum of HCRBs, ECRBs, and MSEs for the estimation of timing offsets or channels from all the nodes are evaluated in each simulation run. In all simulations, the network is assumed to be equipped with $K = 4$ nodes. The oversampling factor is set to $Q = 2$ and a RRC filter with a roll-off factor of 0.22 is employed. The normalized timing offsets from nodes to destination, τ_k, $\forall$ k, are uniformly distributed over the whole symbol period duration, i.e., $\tau_k \sim \mathcal{U}(-0.5, 0.5)$. The channels, h_k, $\forall$ k, are modeled as independent and identically distributed complex Gaussian random variables with $\mathcal{CN}(0, \sigma_h^2 = 1)$.

A. Estimation Performance of Different TSs

Fig. 6 plots the HCRB and ECRB for MTOs estimation versus SNR for the TSs in Table II. Since the HCRB for

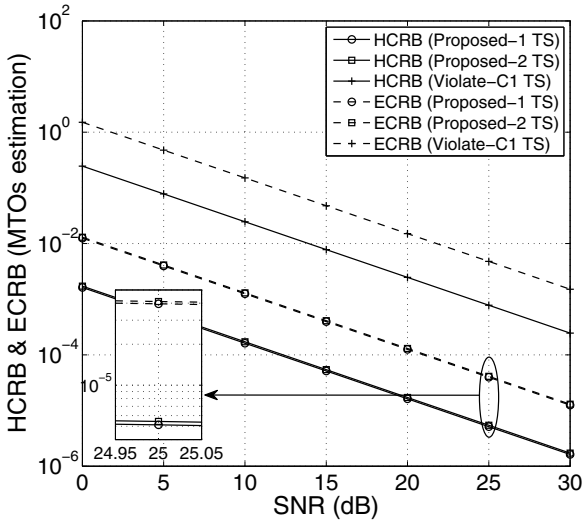


Fig. 6. HCRB and ECRB versus SNR for the estimation of $\boldsymbol{\tau}$ for the TSs given in Table II.

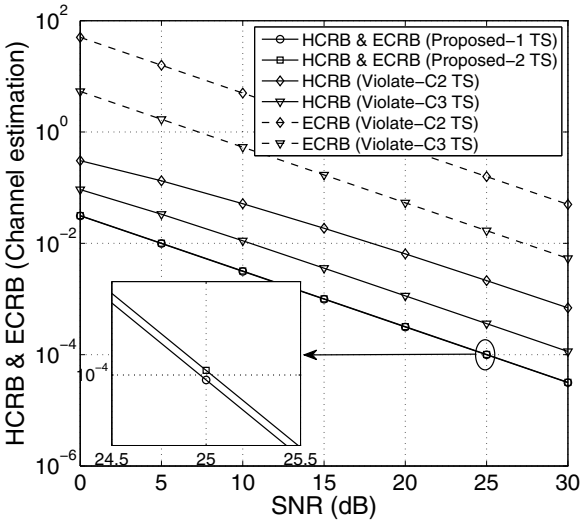


Fig. 7. HCRB and ECRB versus SNR for the estimation of $\mathbf{h}$ for the TSs given in Table II.

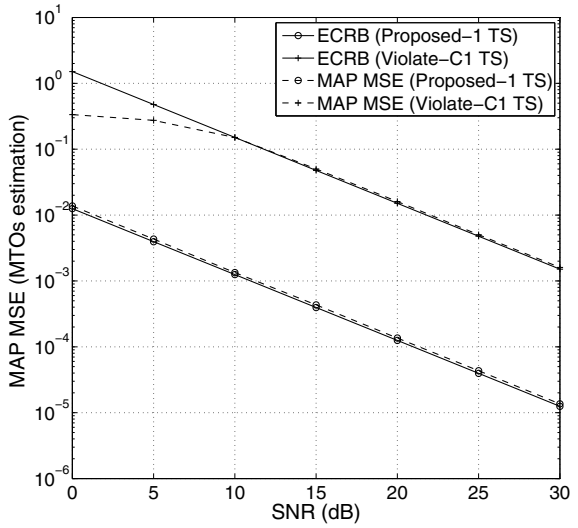


Fig. 8. MSE of MAP estimator and ECRB versus SNR for the estimation of $\boldsymbol{\tau}$ for the TSs given in Table II.

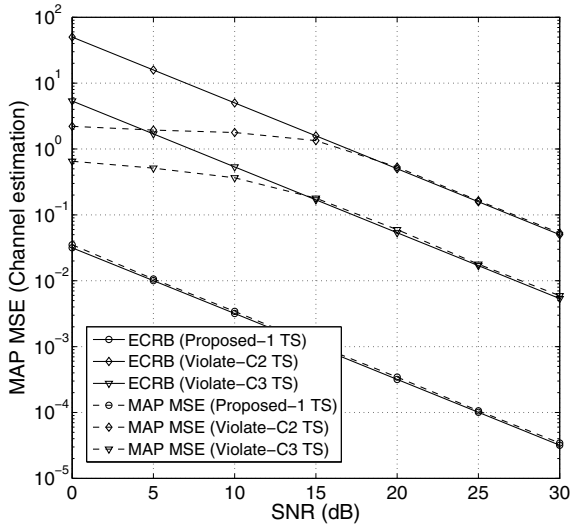


Fig. 9. MSE of MAP estimator and ECRB versus SNR for the estimation of $\mathbf{h}$ for the TSs given in Table II.

MTO estimation is not affected by violating C2 and C3 (see Section IV-A), the results for TSs Violate-C2 and Violate-C3 are not plotted. Fig. 6 demonstrates that the HCRB and ECRB for MTO estimation are the lowest for the Proposed-1 and Proposed-2 TSs and worst for the Violate-C1 TS.

Fig. 7 plots the HCRB and ECRB for multiple channel estimation versus SNR for the Proposed-1, Proposed-2, Violate-C2, and Violate-C3 TSs. The results for the Violate-C1 TS is not plotted since the HCRB for multiple channel estimation is only affected by violating C2 and C3 (see Section IV-B). It can be observed from Fig. 7 that the HCRB and ECRB for multiple channel estimation are the lowest for the TSs, Proposed-1 and Proposed-2, and worst for the TSs Violate-C2 and Violate-C3. Further, numerical results show that the HCRB and ECRB for the Proposed-1 and Proposed-2 TSs are very close. Thus, to avoid repetition, only a single curve is plotted for both the HCRB and ECRB of the Proposed-1 and Proposed-2 TSs. Finally, as mentioned in Section III-B, Fig. 6 and Fig. 7 show that the TSs that minimize the HCRB also minimize the ECRB.

Fig. 8 plots the MAP estimator's MSE for MTOs estimation versus SNR for the Proposed-1 and Violate-C1 TSs. For clarity, the results for the Proposed-2 TS are not plotted since they are very close to that of the Proposed-1 TS. It can be observed from Fig. 8 that being based on exhaustive search, the MAP estimator's MSE is close to the ECRB over a wide range of SNR values for both the Proposed-1 and Violate-C1 TSs. Fig. 8 also shows that for the Violate-C1 TS, the MSE of the MAP estimator is lower than the ECRB at low SNR since the MAP estimator's estimation range is limited to $(-0.5, 0.5)$, given that the timing offset values, τ_k, are assumed to be $\tau_k \in (-0.5, 0.5)$, $\forall k$. However, in its inherent structure, ECRB does not take the range of possible timing offset values into account and grows without bound as the

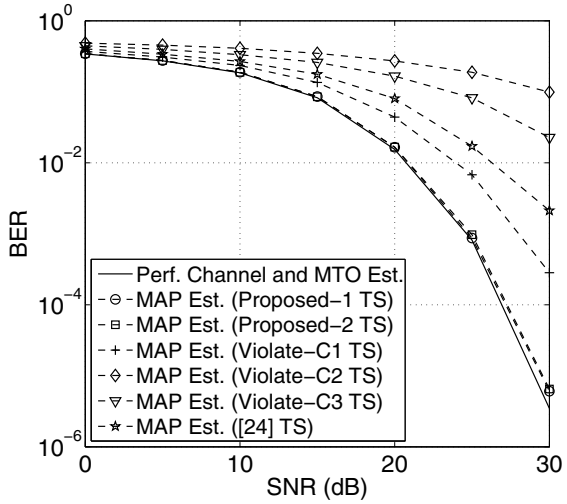


Fig. 10. BER performance of overall system versus SNR for the TSs given in Table II and existing TS in [24], with TS length = 64, data transmission length = 576, 256-QAM modulation, and synchronization overhead = 10%.

SNR decreases (for more information refer to [25] and [36]).

Fig. 9 plots the MAP estimator's MSE for multiple channel estimation versus SNR for the Proposed-1, Violate-C2, and Violate-C3 TSs. It can be observed from Fig. 9 that the MSE performance of the MAP estimator is close to the ECRB at moderate-to-high SNR values for all TSs. Similar to Fig. 8, the MAP estimator's MSE and the ECRB for the Proposed-2 TS are not plotted since the results are similar to that of Proposed-1 TS. Moreover, since the MAP estimator takes the prior information on the distribution of channels and the range of timing offset values into account (see (20)) and the ECRB does not (see (6)), the estimator's MSE is lower than the ECRB at low SNR for the Violate-C2 and Violate-C3 TSs.

From Figs. 6–9 it can be concluded that the Proposed-1 and Proposed-2 TSs can significantly improve estimation accuracy since they result in the lowest HCRB, ECRB, and MAP estimation MSE. These results also demonstrate the large performance gain in terms of estimation accuracy for the proposed TSs over non-optimal TSs that violate the proposed conditions, which numerically validates the proposed conditions for the design of optimal TSs in Section IV-C.

B. System BER performance of different TSs

Fig. 10 shows the end-to-end BER performance of a DF 4-relay cooperative network versus SNR for all the TSs given in Table II. The proposed MAP estimator is applied for joint estimation of MTOs and multiple channels during the training period. 256-quadrature amplitude modulation (256-QAM) is employed for data transmission and length of the source data vector is set to 576 symbols (frame length $= 576 + 64 = 640$ symbols) during the data transmission period, resulting in a synchronization overhead of 10%. Distributed space time block codes are applied at all relays to exploit spatial diversity [37], [38]. To decode the source signal $\mathbf{s}$, a minimum mean-square error linear receiver given by

$$\hat{\mathbf{s}} = (\mathbf{\Lambda}^H \mathbf{\Lambda} + \sigma_w^2 \mathbf{I}_{L\times L})^{-1}\mathbf{y},$$

is employed at the destination node, where $\mathbf{\Lambda} \triangleq \sum_{k=1}^{K} \hat{h}_k \hat{\mathbf{G}}_k$, $\hat{\mathbf{G}}_k = \mathbf{G}_k|_{\tau_k=\hat{\tau_k}}$. The BER performance of the overall system using different TSs is also compared with the benchmark BER, which assumes perfect knowledge of MTO and channels. Fig. 10 shows that for the Proposed-1 and Proposed-2 TSs, the BER performance of the overall system is very close to the benchmark BER plot over a wide range of SNR values, i.e., performance gap of only 0.4 dB lies between the benchmark and proposed system's BER at moderate-to-high SNR values. Fig. 10 also compares the BER performance of the proposed TSs with the TSs proposed in [24]. Fig. 10 shows that our proposed TSs significantly outperform the TS in [24]. Similarly, for the Violate-C1, Violate-C2, and Violate-C3 TSs, the BER results show poor performance. Specifically, for the TSs, Violate-C1 and Violate-C2, receiver almost fails to decode the received signal at the destination for SNR $<$ 30 dB. It has been found through simulations that for $K = 2$ users, satisfying C1 is more critical than satisfying C2 and C3, while for $K > 2$ users, satisfying C2 is more crucial than satisfying C1 and C3.[4] Fig. 10 demonstrates that large BER performance gain can be achieved by employing the proposed TSs compared to non-optimal TSs that violate the proposed conditions C1-C3.

Finally, we have observed through simulations that increasing the TS length for a given fixed frame length, improves the BER performance of the overall system and the results asymptotically converge to the BER with perfect synchronization. Thus, the TS length needs to be carefully selected by system designers to achieve the desired system performance at a specific SNR.

VII. Conclusions

In this paper, the optimal TS design for efficient and joint estimation of MTOs and multiple channels in distributed multi-user and multi-relay cooperative networks is addressed. The HCRB, ECRB, and MAP estimator for estimation of parameters of interest are derived. Next, using these results, guidelines for the design of TSs that minimize the HCRB for joint estimation of timing offsets and channels are proposed. It has been observed through numerical simulations that the ECRB serves as a tighter lower bound for the MAP estimator's MSE at moderate-to-high SNR. Our proposed training guidelines show that the optimal TSs that jointly minimize the HCRB of MTO and multiple channel estimation satisfy three conditions: C1. the optimal TSs from the different nodes exhibit π radian phase shift every symbol period, C2. they are mutually orthogonal, and C3. they are orthogonal to $\pm$ T-shifted TSs from every other node. Numerical results show that the proposed optimal TS design conditions not only lower the HCRB, but also lower the tighter bound, ECRB and the MSE of the derived MAP estimator. Moreover, by applying the proposed guidelines, two TSs are proposed and the estimation MSE and BER performance of the proposed TSs are compared against non-optimal TSs. Simulation results demonstrated large performance gain in terms of estimation

[4]This is observed through simulations but the results for $K = 2$ and 3 users are not included here due to space limitations.

accuracy and end-to-end BER performance when applying the proposed TSs compared to other TS choices.

The proposed TSs, which are summarized in Table II, can also be used for other network topologies, e.g., multi-hop systems and star networks. In multi-hop systems, the proposed TSs can be applied to estimate MTOs and channel parameters in each hop. In star networks, there exist point-to-point communication links between all nodes and the proposed TS with $K = 1$ can be applied to estimate the timing offset and channel parameters corresponding each link. The design of optimal TSs for joint MTO, MCFO, and channel estimation in cooperative systems is subject of future research.

APPENDIX A
DERIVATION OF HIM

In this appendix, a closed-form expression for the **HIM** in (3) is derived. The **HIM** for the estimation of parameters of interest, $\boldsymbol{\theta}$, given the observation vector, $\mathbf{y}$, is given by [31, page 12]

$$\mathbf{HIM} = \mathbb{E}_{\boldsymbol{\theta}_r|\boldsymbol{\theta}_d}[\mathbf{FIM}] + \mathbf{PIM}, \tag{A.1}$$

where **FIM** is the Fisher's information matrix for the estimation of $\boldsymbol{\theta}$ and **PIM** is the prior information matrix (PIM), such that

$$\mathbb{E}_{\boldsymbol{\theta}_r|\boldsymbol{\theta}_d}[\mathbf{FIM}] = \frac{2}{\sigma_w^2}\begin{bmatrix} \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & -\Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \mathbf{0}_{K\times K} \\ \Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \mathbf{0}_{K\times K} \\ \mathbf{0}_{K\times K} & \mathbf{0}_{K\times K} & \Re\{\mathbf{U}\} \end{bmatrix},$$
$$\mathbf{PIM} = \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{\theta}_r}^{-1} & \mathbf{0}_{2K\times K} \\ \mathbf{0}_{K\times 2K} & \mathbf{0}_{K\times K} \end{bmatrix}. \tag{A.2}$$

Since $\boldsymbol{\theta}_r$ and $\boldsymbol{\theta}_d$, defined below (2), are independent parameters, $p(\boldsymbol{\theta}_r|\boldsymbol{\theta}_d) = p(\boldsymbol{\theta}_r)$ and $\mathbb{E}_{\boldsymbol{\theta}_r|\boldsymbol{\theta}_d}[\mathbf{FIM}] = \mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{FIM}]$ in (A.2) [31, page 12]. The detailed derivation of $\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{FIM}]$ and **PIM** is given in the following subsections.

A. Derivation of $\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{FIM}]$

FIM for joint estimation of MTOs and channel gains is given by (A.3) at the bottom of this page [25], where $\mathbf{D} \triangleq \text{diag}(h_1, \ldots, h_K)$ is a $K \times K$ diagonal matrix and $\boldsymbol{\Delta} \triangleq \frac{\partial \boldsymbol{\Psi}}{\partial \boldsymbol{\tau}} \triangleq [\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_K]$ is a $QL \times K$ matrix. To find $\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{FIM}]$, we have to find the expected value of all the elements of **FIM** w.r.t. $\boldsymbol{\theta}_r$. Using the fact that $\Re\{\cdot\}$ and $\Im\{\cdot\}$ are linear operators, $\mathbb{E}[\Re\{\cdot\}] = \Re\{\mathbb{E}[\cdot]\}$ and $\mathbb{E}[\Im\{\cdot\}] = \Im\{\mathbb{E}[\cdot]\}$. Thus $\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{FIM}]$ is given in (A.4) at the bottom of this page. In order to calculate the individual elements of $\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{FIM}]$, we have to find the distribution of $\boldsymbol{\theta}_r$. The channels from different nodes to the receiver are modeled as independent and identically distributed random variables, i.e., $h_k \sim \mathcal{CN}(0, \sigma_h^2)$, $\forall k$. Thus, $\boldsymbol{\theta}_r$ is multivariate normal distributed with mean zero and covariance $\boldsymbol{\Sigma}_{\boldsymbol{\theta}_r}$, i.e., $\boldsymbol{\theta}_r \sim \mathcal{N}(\mathbf{0}_{2K\times 1}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}_r})$, where $\boldsymbol{\Sigma}_{\boldsymbol{\theta}_r} = \text{diag}\Big(\underbrace{\frac{\sigma_h^2}{2}, \ldots, \frac{\sigma_h^2}{2}}_{2K}\Big)$. Thus, the PDF of $\boldsymbol{\theta}_r$, $p(\boldsymbol{\theta}_r)$, is given by

$$p(\boldsymbol{\theta}_r) = \frac{1}{(2\pi)^K \det(\boldsymbol{\Sigma}_{\boldsymbol{\theta}_r})^{1/2}} \exp\left\{-\frac{\boldsymbol{\theta}_r^T \boldsymbol{\Sigma}_{\boldsymbol{\theta}_r}^{-1} \boldsymbol{\theta}_r}{2}\right\}. \tag{A.5}$$

Using (A.5), the submatrices of $\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{FIM}]$ can be determined. Thus, $\mathbb{E}_{\boldsymbol{\theta}_r}[\boldsymbol{\Psi}^H\boldsymbol{\Delta}\mathbf{D}]$ is given by

$$\mathbb{E}_{\boldsymbol{\theta}_r}[\boldsymbol{\Psi}^H\boldsymbol{\Delta}\mathbf{D}] = \begin{bmatrix} \mathbb{E}_{\boldsymbol{\theta}_r}[h_1\boldsymbol{\xi}_1^H\boldsymbol{\delta}_1] & \ldots & \mathbb{E}_{\boldsymbol{\theta}_r}[h_K\boldsymbol{\xi}_1^H\boldsymbol{\delta}_K] \\ \vdots & \ddots & \vdots \\ \mathbb{E}_{\boldsymbol{\theta}_r}[h_1\boldsymbol{\xi}_K^H\boldsymbol{\delta}_1] & \ldots & \mathbb{E}_{\boldsymbol{\theta}_r}[h_K\boldsymbol{\xi}_K^H\boldsymbol{\delta}_K] \end{bmatrix} = \mathbf{0}_{K\times K}. \tag{A.6}$$

The equality in (A.6) follows from the fact the channel gains, h_k, $\forall\ k$, are zero-mean random variables. It can be similarly concluded that in (A.4)

$$\mathbb{E}_{\boldsymbol{\theta}_r}[\boldsymbol{\Psi}^H\boldsymbol{\Delta}\mathbf{D}] = \mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Psi}] = \mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Psi}] = \mathbf{0}_{K\times K}. \tag{A.7}$$

Finally, $\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Delta}\mathbf{D}]$ is given by (A.8) at the bottom of this page, where the off-diagonal elements of (A.8) are zero since the channels are uncorrelated, i.e., $\mathbb{E}_{\boldsymbol{\theta}_r}[h_k^* h_{\bar{k}}] = 0$ for $k, \bar{k} = 1, \ldots, K$, and $k \neq \bar{k}$. The diagonal elements of (A.8) for $k = 1, \ldots, K$ are given by

$$\mathbb{E}_{\boldsymbol{\theta}_r}[|h_k|^2\boldsymbol{\delta}_k^H\boldsymbol{\delta}_k] = \boldsymbol{\delta}_k^H\boldsymbol{\delta}_k\Big(\mathbb{E}_{\Re\{h_{\bar{k}}\}}\big[\Re\{h_{\bar{k}}\}^2\big] + \mathbb{E}_{\Im\{h_{\bar{k}}\}}\big[\Im\{h_{\bar{k}}\}^2\big]\Big) = \sigma_h^2\boldsymbol{\delta}_k^H\boldsymbol{\delta}_k. \tag{A.9}$$

Substituting (A.6), (A.7), and (A.8) into (A.4), the final result in (A.2) can be obtained.

$$\mathbf{FIM} = \frac{2}{\sigma_w^2}\begin{bmatrix} \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & -\Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Delta}\mathbf{D}\} \\ \Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \Im\{\boldsymbol{\Psi}^H\boldsymbol{\Delta}\mathbf{D}\} \\ \Re\{\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Psi}\} & -\Im\{\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Psi}\} & \Re\{\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Delta}\mathbf{D}\} \end{bmatrix}. \tag{A.3}$$

$$\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{FIM}] = \frac{2}{\sigma_w^2}\begin{bmatrix} \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & -\Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \Re\{\mathbb{E}_{\boldsymbol{\theta}_r}[\boldsymbol{\Psi}^H\boldsymbol{\Delta}\mathbf{D}]\} \\ \Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \Im\{\mathbb{E}_{\boldsymbol{\theta}_r}[\boldsymbol{\Psi}^H\boldsymbol{\Delta}\mathbf{D}]\} \\ \Re\{\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Psi}]\} & -\Im\{\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Psi}]\} & \Re\{\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Delta}\mathbf{D}]\} \end{bmatrix}. \tag{A.4}$$

$$\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Delta}\mathbf{D}] = \begin{bmatrix} \mathbb{E}_{\boldsymbol{\theta}_r}[|h_1|^2\boldsymbol{\delta}_1^H\boldsymbol{\delta}_1] & \mathbb{E}_{\boldsymbol{\theta}_r}[h_1^*\boldsymbol{\delta}_1^H\boldsymbol{\delta}_2 h_2] & \ldots & \mathbb{E}_{\boldsymbol{\theta}_r}[h_1^*\boldsymbol{\delta}_1^H\boldsymbol{\delta}_K h_K] \\ \mathbb{E}_{\boldsymbol{\theta}_r}[h_2^*\boldsymbol{\delta}_2^H\boldsymbol{\delta}_1 h_1] & \mathbb{E}_{\boldsymbol{\theta}_r}[|h_2|^2\boldsymbol{\delta}_2^H\boldsymbol{\delta}_2] & \ldots & \mathbb{E}_{\boldsymbol{\theta}_r}[h_2^*\boldsymbol{\delta}_2^H\boldsymbol{\delta}_K h_K] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}_{\boldsymbol{\theta}_r}[h_K^*\boldsymbol{\delta}_K^H\boldsymbol{\delta}_1 h_1] & \mathbb{E}_{\boldsymbol{\theta}_r}[h_K^*\boldsymbol{\delta}_K^H\boldsymbol{\delta}_2 h_2] & \ldots & \mathbb{E}_{\boldsymbol{\theta}_r}[|h_K|^2\boldsymbol{\delta}_K^H\boldsymbol{\delta}_K] \end{bmatrix} = \sigma_h^2 \text{diag}\left(\boldsymbol{\delta}_1^H\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_K^H\boldsymbol{\delta}_K\right). \tag{A.8}$$

B. Derivation of $\mathbf{PIM}$

Given the fact that $\mathbf{h}$ and $\boldsymbol{\tau}$ are independent and using the definition of PIM in [31, page 12] and [32, eq. (3)], the $\mathbf{PIM}$ for random channels and deterministic timing offsets can be written as

$$\mathbf{PIM} = \begin{bmatrix} \mathbb{E}_{\boldsymbol{\theta}_r}\big[-\Delta^{\boldsymbol{\theta}_r}_{\boldsymbol{\theta}_r}\log p(\boldsymbol{\theta}_r)\big] & \mathbf{0}_{2K\times K} \\ \mathbf{0}_{K\times 2K} & \mathbf{0}_{K\times K} \end{bmatrix}, \quad \text{(A.10)}$$

where $p(\boldsymbol{\theta}_r)$ is given in (A.5). The negative Hessian of the log likelihood function of $\boldsymbol{\theta}_r$, $-\Delta^{\boldsymbol{\theta}_r}_{\boldsymbol{\theta}_r}\log p(\boldsymbol{\theta}_r)$ is given by

$$-\Delta^{\boldsymbol{\theta}_r}_{\boldsymbol{\theta}_r}\log p(\boldsymbol{\theta}_r) = \frac{1}{2}\Delta^{\boldsymbol{\theta}_r}_{\boldsymbol{\theta}_r}\big({\boldsymbol{\theta}_r}^T\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\theta}_r}\boldsymbol{\theta}_r\big) = \boldsymbol{\Sigma}^{-1}_{\boldsymbol{\theta}_r}. \quad \text{(A.11)}$$

By using the result in (A.11) in (A.10), the final result in (A.2) follows.

APPENDIX B
POSITIVE DEFINITENESS OF $\boldsymbol{\Psi}^H\boldsymbol{\Psi}$

In order to show that $\boldsymbol{\Psi}^H\boldsymbol{\Psi}$ is a positive definite matrix, it is required to show that for any $K\times 1$ vector $\mathbf{x}$, $\mathbf{x}^H\boldsymbol{\Psi}^H\boldsymbol{\Psi}\mathbf{x} > 0$. Clearly

$$\mathbf{x}^H\boldsymbol{\Psi}^H\boldsymbol{\Psi}\mathbf{x} = \|\boldsymbol{\Psi}\mathbf{x}\|^2 \geq 0, \quad \text{(B.1)}$$

which demonstrates that $\boldsymbol{\Psi}^H\boldsymbol{\Psi}$ is a positive *semidefinite* matrix and its eigenvalues are nonnegative [39]. It is known that if a matrix $\boldsymbol{\Psi}^H\boldsymbol{\Psi}$ is full rank, no eigenvalue of $\boldsymbol{\Psi}^H\boldsymbol{\Psi}$ is equal to zero [39]. Since linearly independent TSs are transmitted from all the nodes (see Section III-A), it can be concluded that $\boldsymbol{\Psi}^H\boldsymbol{\Psi}$ is a full rank matrix. Thus, using (B.1) and the full rank nature of $\boldsymbol{\Psi}^H\boldsymbol{\Psi}$, it can be concluded that all eigenvalues of $\boldsymbol{\Psi}^H\boldsymbol{\Psi}$ are strictly positive and $\boldsymbol{\Psi}^H\boldsymbol{\Psi}$ is a positive definite matrix.

REFERENCES

[1] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers, Synchronization, Channel Estimation, and Signal Processing*, J. G. Proakis, Ed. Wiley Series in Telecommunications and Signal Processing, 1998.

[2] Y.-C. Wu and E. Serpedin, "Training sequences design for symbol timing estimation in MIMO correlated fading channels," in *Proc. 2004 IEEE Global Commun. Conf.*

[3] A. Kannan, T. P. Krauss, and M. D. Zoltowski, "Separation of cochannel signals under imperfect timing and carrier synchronization," *IEEE Trans. Veh. Technol.*, vol. 50, no. 1, pp. 79–96, Jan. 2001.

[4] A. A. Nasir, H. Mehrpouyan, S. D. Blostein, S. Durrani, and R. A. Kennedy, "Timing and carrier synchronization with channel estimation in multi-relay cooperative networks," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 793–811, Feb. 2012.

[5] M. Dong and L. Tong, "Optimal design and placement of pilot symbols for channel estimation," *IEEE Trans. Signal Process.*, vol. 50, no. 12, pp. 3055–3069, Dec. 2002.

[6] L. Yang, X. Ma, and G. B. Giannakis, "Optimal training for MIMO fading channels with time- and frequency-selectivity," in *Proc. 2004 IEEE ICASSP*.

[7] S.-A. Yang and J. Wu, "Optimal binary training sequence design for multiple-antenna systems over dispersive fading channels," *IEEE Trans. Veh. Technol.*, vol. 51, no. 5, pp. 1271–1276, Sep. 2002.

[8] F. Gao, T. Cui, and A. Nallanathan, "On channel estimation and optimal training design for amplify-and-forward relay networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1907–1916, May 2008.

[9] F. Gao, R. Zhang, and Y.-C. Liang, "Optimal channel estimation and training design for two-way relay networks," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 3024–3033, Oct. 2009.

[10] Y. Chi, A. Gomaa, N. Al-Dhahir, and R. Calderbank, "Training signal design and tradeoffs for spectrally-efficient multi-user MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 2234–2245, 2011.

[11] E. Bjornson and B. Ottersten, "A framework for training-based estimation in arbitrarily correlated rician MIMO channels with Rician disturbance," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1807–1820, Mar. 2010.

[12] T. Kong and Y. Hua, "Optimal design of source and relay pilots for MIMO relay channel estimation," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4438–4446, Sep. 2011.

[13] X. Zhou, P. Sadeghi, T. A. Lamahewa, and S. Durrani, "Design guidelines for training-based MIMO systems with feedback," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 4014–4026, Oct. 2009.

[14] C. Tellambura, M. Parker, Y. J. Guo, S. Sheperd, and S. Barton, "Optimal sequence for channel estimation using discrete Fourier transform techniques," *IEEE Trans. Commun.*, vol. 47, pp. 230–238, Feb. 1999.

[15] W. Chen and U. Mirta, "Training sequence optimization: comparisons and an alternative criterion," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 1987–1991, Dec. 2000.

[16] S. Zhang, F. Gao, and C.-X. Pei, "Optimal training design for individual channel estimation in two-way relay networks," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4987–4991, Sep. 2012.

[17] C. Fragouli, N. Al-Dhahir, and W. Turin, "Training-based channel estimation for multiple-antenna broadband transmissions," *IEEE Trans. Wireless Commun.*, vol. 2, no. 2, pp. 384–391, Mar. 2003.

[18] F. Gao, T. Cui, and A. Nallanathan, "Optimal training design for channel estimation in decode-and-forward relay networks with individual and total power constraints," *IEEE Trans. Signal Processing*, vol. 56, no. 12, pp. 5937–5949, Dec. 2008.

[19] T. Cui and C. Tellambura, "Joint channel and frequency offset estimation and training sequence design for MIMO systems over frequency selective channels," in *Proc. 2004 IEEE Global Commun. Conf.*

[20] F. Simoens and M. Moeneclaey, "Computationally efficient frequency offset estimation for flat-fading MIMO channels: performance analysis and training sequence design," in *Proc. 2004 IEEE Global Commun. Conf.*

[21] M. Ghogho and A. Swami, "Training design for channel and CFO estimation in MIMO systems," in *Proc. 2006 IEEE ICASSP*.

[22] P. A. Parker, P. Mitran, D. W. Bliss, and V. Tarokh, "Training sequence design for wireless collaborative communication systems in frequency-selective fading," in *Proc. 2007 IEEE Asilomar Conf. Signals, Systems, Computers*.

[23] P. Stoica and O. Besson, "Training sequence design for frequency offset and frequency-selective channel estimation," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1910–1917, Nov. 2003.

[24] H. Mehrpouyan and S. D. Blostein, "Estimation, training, and effect of timing offsets in distributed cooperative networks," in *Proc. 2010 IEEE Global Commun. Conf.*

[25] X. Li, Y. C. Wu, and E. Serpedin, "Timing synchronization in decode-and-forward cooperative communication systems," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1444–1455, Apr. 2009.

[26] X. Li, C. Xing, Y.-C. Wu, and S. C. Chan, "Timing estimation and resynchronization for amplify-and-forward communication systems," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2218–2229, Apr. 2010.

[27] Y.-C. Wu, S. C. Chan, and E. Serpedin, "Symbol-timing estimation in spacetime coding systems based on orthogonal training sequences," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 603–613, Mar. 2005.

[28] A. F. Naguib, V. Tarokh, N. Seshadri, and A. R. Calderbank, "A spacetime coding modem for high-data-rate wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 1459–1478, Oct. 1998.

[29] K. Rajawat and A. K. Chaturvedi, "A low complexity symbol timing estimator for MIMO systems using two samples per symbol," *IEEE Commun. Lett.*, vol. 10, no. 7, pp. 525–527, Jul. 2006.

[30] P. A. Parker, P. Mitran, D. W. Bliss, and V. Tarokh, "On bounds and algorithms for frequency synchronization for collaborative communication systems," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3742–3752, Aug. 2008.

[31] H. L. V. Trees and K. L. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering / Tracking*, M. E. El-Hawary, Ed. Wiley interscience, 2007.

[32] S. Bay, B. Geller, A. Renaux, J.-P. Barbot, and J.-M. Brossier, "On the hybrid cramr rao bound and its application to dynamical phase estimation," *IEEE Signal Process. Lett.*, vol. 15, pp. 453–456, May 2008.

[33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[34] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.

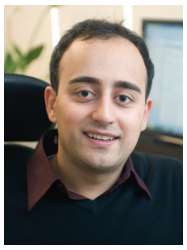
[35] G. L. Stuber, *Principles of Mobile Communication*. Springer, 2011.

[36] A. A. Nasir, S. Durrani, and R. A. Kennedy, "Particle filters for joint timing and carrier estimation: improved resampling guidelines

and weighted Bayesian Cramer-Rao bounds," *IEEE Trans. Commun.*, vol. 60, no. 5, pp. 1407–1419, May 2012.
[37] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3524–3536, Dec. 2006.
[38] Y. Jing and H. Jafarkhani, "Using orthogonal and quasi-orthogonal designs in wireless relay networks," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4106–4118, Nov. 2007.
[39] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1990.

Ali A. Nasir received his B.Sc. (1st class hons.) degree in Electrical Engineering from the University of Engineering and Technology (UET), Lahore, Pakistan in 2007. He joined the Research School of Engineering at the Australian National University (ANU), Australia in 2009, where he received his Ph.D. in Electrical Engineering in 2013. He is currently a Post Doctoral research fellow in the Applied Signal Processing group at ANU. Prior to joining ANU, he worked as a Lecturer in UET for nine months and then as a Design Engineer in the Centre for Advanced Research in Engineering (CARE), Islamabad, Pakistan for a year. He was a research visitor in Chalmers University of Technology, Gothenburg, Sweden for six months from April till October of 2011.

Ali was awarded a University Gold Medal for outstanding performance during the final year of his undergraduate studies. He is a recipient of an ANU International Ph.D. scholarship for the duration of his Ph.D. He was also awarded an ANU Vice Chancellor's Higher Degree Research (HDR) travel grant in 2011. He is an Associate Editor for *IEEE Canadian Journal of Electrical and Computer Engineering*. His research interests are in the area of synchronization and channel estimation in cooperative communication, MIMO and OFDM systems.



Hani Mehrpouyan (S'05-M'10) received his B.Sc. honours degree in Computer Engineering from Simon Fraser University, Burnaby, Canada in 2004. In 2005 he joined the Department of ECE at Queens University, Kingston, Canada as a Ph.D. candidate, where he received his Ph.D. in Electrical Engineering in 2010. From September 2010 to March of 2012 he was a Post-Doc at the Department of Signal and Systems at Chalmers University of Technology where he lead the MIMO aspects of the microwave backhauling for next generation wireless networks project. He next joined the University of Luxembourg as a Research Associate from April 2012 to August of 2012, where he was responsible for new interference cancelation and synchronization schemes for next generation satellite communication links. Since August of 2012 he has been an Assistant Professor at the Department of Computer and Electrical Engineering and Computer Science at California State University, Bakersfield.

Dr. Mehrpouyan has received more than 10 scholarships and awards, e.g., IEEE Wireless Communication exemplary reviewer award, IEEE Globecom Early Bird Student Award, NSERC-IRDF, NSERC PGS-D, NSERC CGS-M Alexander Graham Bell, B.C. Wireless Innovation, and more. He has more than 30 publications in prestigious IEEE Journals and Conferences. He has served as a TPC member for IEEE Globecom 2012. He has also been service as the associate editor for IEEE Canadian Journal of Electrical Engineering. Dr Mehrpouyan has also been involved with industry leaders such as Ericsson AB, Research in Motion (RIM), Alcatel etc. His current research interests lie in the area of applied signal processing and physical layer of wireless communication systems, including synchronization, channel estimation, interference cancelation, and performance optimization. For more information refer to www.mehrpouyan.info.

Salman Durrani (S'00-M'05-SM'10) received the B.Sc. (1st class honours) degree in Electrical Engineering from the University of Engineering & Technology, Lahore, Pakistan in 2000. He received the Ph.D. degree in Electrical Engineering from the University of Queensland, Brisbane, Australia in Dec. 2004. He has been with the Australian National University, Canberra, Australia, since 2005, where he is currently Senior Lecturer in the Research School of Engineering, College of Engineering & Computer Science.

Dr. Durrani serves as a Technical Program Committee Member for international conferences such as ICC '13, ICSPCS '13 and PIMRC '12. He was awarded an ANU Vice-Chancellor's Award for Teaching Excellence in 2012. He was a recipient of an International Postgraduate Research Scholarship from the Australian Commonwealth during 2001-2004. He was awarded a University Gold Medal during his undergraduate studies. His current research interests are in wireless communications and signal processing, including synchronization in cooperative communication systems, interference, outage and connectivity of wireless networks and signal processing on the unit sphere. He has over 60 publications to date in refereed international journals and conferences. He is a Member of Institution of Engineers, Australia.

Steven D. Blostein (S'83-M'88-SM'96) received his B.S. degree in Electrical Engineering from Cornell University, Ithaca, NY, in 1983, and the M.S. and Ph.D. degrees in Electrical and Computer Engineering from the University of Illinois, Urbana-Champaign, in 1985 and 1988, respectively. He has been on the faculty in the Department of Electrical and Computer Engineering Queen's University since 1988 and currently holds the position of Professor. From 2004-2009 he was Department Head. From 1999-2003, he was the leader of the Multi-Rate Wireless Data Access Major Project sponsored by the Canadian Institute for Telecommunications Research. He has also been a consultant to industry and government in the areas of image compression, target tracking, radar imaging and wireless communications. He spent sabbatical leaves at Lockheed Martin Electronic Systems and at Communications Research Centre in Ottawa. His current interests lie in the application of signal processing to problems in wireless communications systems, including synchronization, network MIMO and physical layer optimization for multimedia transmission. He has been a member of the Samsung 4G Wireless Forum as well as an invited distinguished speaker. He served as Chair of IEEE Kingston Section (1994), Chair of the Biennial Symposium on Communications (2000,2006,2008), associate editor for IEEE TRANSACTIONS ON IMAGE PROCESSING (1996-2000), and Publications Chair for IEEE ICASSP 2004. For a number of years has been serving on technical program committees for IEEE Communications Society conferences that include ICC, Globecom and WCNC. He has been serving as an editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS since 2007. He is a registered professional engineer in Ontario and a Senior Member of IEEE.

Rodney A. Kennedy S'86-M'88-SM'01-F'05) received the B.E. degree from the University of New South Wales, Sydney, Australia, the M.E. degree from the University of Newcastle, and the Ph.D. degree from the Australian National University, Canberra. He is currently a Professor in the Research School of Engineering, Australian National University. He is a Fellow of the IEEE. His research interests include digital signal processing, digital and wireless communications, and acoustical signal processing.

Björn Ottersten was born in Stockholm, Sweden, 1961. He received the M.S. degree in electrical engineering and applied physics from Linköping University, Linköping, Sweden, in 1986. In 1989 he received the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA. Dr. Ottersten has held research positions at the Department of Electrical Engineering, Linköping University, the Information Systems Laboratory, Stanford University, the Katholieke Universiteit Leuven, Leuven, and the University of Luxembourg. During 96/97 Dr. Ottersten was Director of Research at ArrayComm Inc, a start-up in San Jose, California based on Ottersten's patented technology. He has co-authored journal papers that received the IEEE Signal Processing Society Best Paper Award in 1993, 2001, and 2006 and 3 IEEE conference papers receiving Best Paper Awards. In 1991 he was appointed Professor of Signal Processing at the Royal Institute of Technology (KTH), Stockholm. From 1992 to 2004 he was head of the department for Signals, Sensors, and Systems at KTH and from 2004 to 2008 he was dean of the School of Electrical Engineering at KTH. Currently, Dr. Ottersten is Director for the Interdisciplinary Centre for Security, Reliability and Trust at the University of Luxembourg. As Digital Champion of Luxembourg, he acts as an adviser to European Commissioner Neelie Kroes. Dr. Ottersten has served as Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and on the editorial board of *IEEE Signal Processing Magazine*. He is currently editor in chief of *EURASIP Signal Processing Journal* and a member of the editorial board of *EURASIP Journal of Applied Signal Processing*. Dr. Ottersten is a Fellow of the IEEE and EURASIP. In 2011 he received the IEEE Signal Processing Society Technical Achievement Award. He is a first recipient of the European Research Council advanced research grant. His research interests include security and trust, reliable wireless communications, and statistical signal processing.