



Title	Timing synchronization in decode-and-forward cooperative communication systems
Author(s)	Li, X; Wu, YC; Serpedin, E
Citation	IEEE Transactions On Signal Processing, 2009, v. 57 n. 4, p. 1444-1455
Issued Date	2009
URL	http://hdl.handle.net/10722/58742
Rights	©2009 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Timing Synchronization in Decode-and-Forward Cooperative Communication Systems

Xiao Li, Yik-Chung Wu, and Erchin Serpedin

Abstract—Cooperative communication systems have attracted much attention recently due to their desirable performance gain while using single antenna terminals. This paper addresses the joint timing and channel estimation problem, and furthermore the resynchronization of multiple timing offsets in a cooperative relay system. The estimations of timing and channel are conducted in two phases and the associated Cramér-Rao bounds (CRB) are derived for both phases. It is demonstrated that the conventional CRB is not valid for timing parameters under fading conditions, and a new bound called Weighted Bayesian CRB is proposed. With the timing and channel estimates, a general framework of the resynchronization filter design is developed in order to compensate the multiple timing offsets at the destination. The proposed methods are applied to different scenarios with varying degrees of timing misalignment and are numerically shown to provide excellent performances that approach the perfectly synchronized case.

Index Terms—Cooperative, timing offset, channel estimation, Cramér-Rao bound (CRB), resynchronization, Tikhonov regularization.

I. INTRODUCTION

MULTIPLE-INPUT-MULTIPLE-OUTPUT (MIMO) systems have been demonstrated to be one of the most promising and potential candidates for the future wireless communications [1]–[3]. However, limited by the cost and the size of equipment, the concept of *distributed* MIMO system has been advocated by many researchers, where the sharing of antennas among several single-antenna terminals [4] to cooperatively transmit data is suggested. It has been pointed out that with proper cooperative strategies, the same benefits of centralized MIMO systems can be achieved in a cooperative MIMO system [5]–[10].

So far, many results regarding cooperative communication systems have been obtained under the assumption of perfect synchronization among the users, which is generally very difficult to achieve due to the distributed nature. Analytical and

numerical results on the performance degradation brought by timing errors in a cooperative system have been reported in [11], [12] and [13]. Furthermore, it is also found that if the timing errors are large, the benefit of cooperation would even vanish in terms of diversity gain [14] and system capacity [15]. All these results demonstrate that appropriate countermeasures to the asynchronous reception in distributed MIMO systems are absolutely necessary.

Some researchers have suggested using delay-robust transmission schemes so that the mentioned problem is bypassed [16]–[20]. While these methods are promising and desirable, they impose restrictions on how data are transmitted and how users are cooperating, thus limiting their applicability. On the other hand, many current cooperative transmission schemes and their demodulation methods require the cooperative users to be temporally synchronized, including (but not limited to) distributed space-time coding, cooperative relays [12], [13], cooperative eigen-coding scheme and distributed unitary space-time modulation (USTM) [21]. Furthermore, frequency synchronization and channel estimation issues in a cooperative system have been addressed by [22] and [23], while timing synchronization issues receive much less attention. Therefore, algorithms that can resynchronize the distributed users at the receiver are of tremendous value.

This paper develops a general framework for the estimation and synchronization problem in cooperative decode-and-forward relay systems. The contributions of this paper are summarized as follows. First, the maximum likelihood (ML) estimators and Cramér-Rao bounds (CRB) are derived for the timing and channel parameters. Second, it is demonstrated that the CRB is not applicable to synchronization problems in the low signal-to-noise ratio (SNR) region and under fading environment. We then derive a Weighted Bayesian CRB (WBCRB), which is a valid bound for all SNRs even in fading channels. Third, resynchronization filters are proposed to compensate the multiple timing offsets at the destination terminal. As a computationally efficient but less accurate method, a matched filter that minimizes the residual intersymbol interference (ISI) power is derived. Furthermore, a more general method to deal with the resynchronization filter design problem is developed by using the weighted regularized least squares (WRLS) approach, with the optimal regularization parameter chosen by the L-curve technique. It is shown that the symbol error rate (SER) after the compensation fundamentally improves and approaches the ideal case when the timing misalignment is relatively mild.

The rest of this paper is organized as follows. In Section II, the system model for the considered relay system is presented.

Manuscript received May 29, 2008; revised November 10, 2008. First published December 31, 2008; current version published March 11, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Mounir Ghogho. The work of X. Li and Y.-C. Wu was supported by Grant GRF HKU 7181/07E.

X. Li and Y.-C. Wu are with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong (e-mail: lixiao@eee.hku.hk; yewu@eee.hku.hk).

E. Serpedin is with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843 USA (e-mail: serpedin@ece.tamu.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSP.2008.2011834

The estimators for both the broadcasting phase and the multiple access phase are investigated and derived in Section III, followed by a thorough CRB analysis in Section IV. The design of the resynchronization filter is discussed in Section V with an in-depth analysis over the proposed methods. Section VI provides numerical results to validate the proposed estimation and resynchronization schemes. Finally, the paper is concluded in Section VII.

Notation : The operators $\Re(\cdot)$ and $\Im(\cdot)$ correspond to the real part and imaginary part of a complex number, while $\text{diag}(\mathbf{x})$ denotes a diagonal matrix with the elements of \mathbf{x} located on the main diagonal. Superscripts $(\cdot)^*$, $(\cdot)^H$, and $(\cdot)^T$ denote the conjugate, the conjugate transpose, and the transpose operators, respectively. Notation \mathbf{I} is the identity matrix, $\|\mathbf{x}\|$ represents the L_2 norm of vector \mathbf{x} , and $\mathbb{E}_\theta\{\cdot\}$ assumes the expectation with respect to variable θ .

II. SYSTEM MODEL

The system under consideration consists of one source, one destination and K clustered-relay terminals located in the middle as shown in Fig. 1. The propagation channels are considered to be quasi-static flat-fading [14]. Furthermore, we only focus on a two hop system because it can be easily extended to a multihop context. The transmission consists of two periods:

- 1) *Training period:* The source transmits a training sequence to the K relays. After the timing offset and channel estimation, each relay adjusts its own clock and sends out K distinct training sequences. At the destination, a joint estimation of the multiple timing offsets and channels is performed. With the estimated timings and channels, a resynchronization filter is designed using the proposed methods.
- 2) *Data transmission period:* After the training period, the source transmits a data sequence to the K relays. With the timing and channel estimates obtained in the previous training period, each relay decodes the incoming message. Afterwards, the K relays cooperatively transmit the decoded data to the destination. At the destination, the incoming signal is convolved with the resynchronization filter to realign the signals and finally the data are decoded. The timing diagram for the two periods is shown in Fig. 2. Within each period, the transmission process is a combination of a single-input-multiple-output (SIMO) broadcasting system and a multiple-input-single-output (MISO) multiple access system.

A. SIMO Broadcasting Phase

Due to the hardware inconsistencies and diverse relay locations, the signals arriving at different relays are not synchronized to each other. The received signal (within $0 \leq t \leq L_o T$) at the k th relay can be expressed as

$$r_k(t) = c_k \sum_{i=-L_g}^{L_o+L_g-1} s(i)g(t-iT-\tau_k T) + w_k(t) \quad (1)$$

where c_k is the complex channel coefficient from the source to the k th relay and is assumed to be zero mean, circular complex

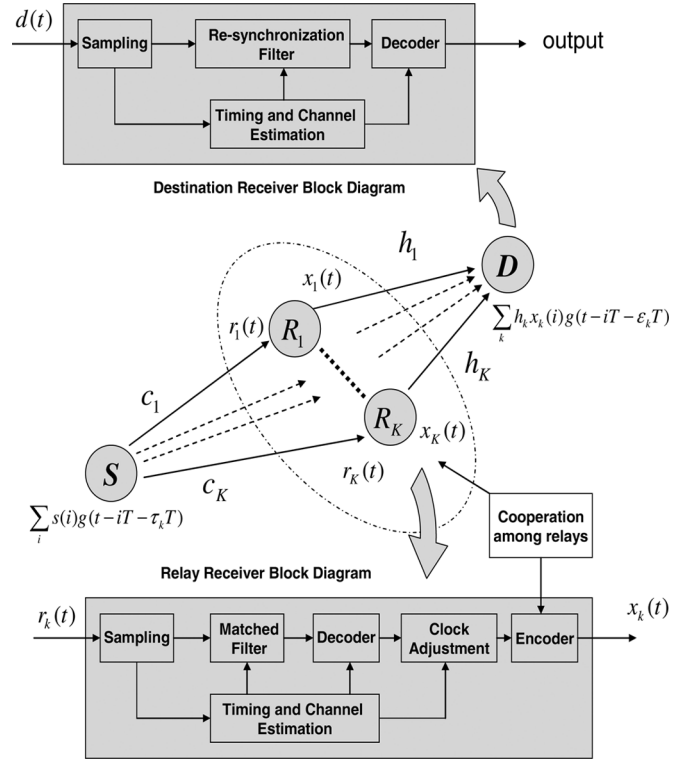


Fig. 1. A typical cooperative relay system.

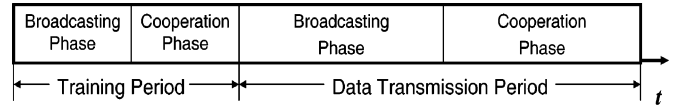


Fig. 2. Timing diagram.

Gaussian random variable with unit variance. The term $w_k(t)$ is the zero mean, circular complex additive white Gaussian noise (AWGN) with variance $\sigma_{w_k}^2$. Notation T is the symbol duration, $s(i)$ is the complex valued symbol with $\mathbb{E}\{|s(i)|^2\} = E_s$; $\tau_k \in [0, 1)$ is the unknown timing offset normalized to the symbol duration, and $g(t)$ is the pulse shaping filter. Symbol L_o represents the observation interval while L_g is the approximated effective duration of the tail of $g(t)$ on one side.

Upon reception, the signal is oversampled at each relay by a ratio $Q \geq 2$ and thus the sample interval is $T_s = T/Q$. By stacking $L_o Q$ received samples, the received vector is given by [24]

$$\mathbf{r}_k = c_k \mathbf{A}_{\tau_k} \mathbf{s} + \mathbf{w}_k \quad (2)$$

where

$$\begin{aligned} \mathbf{r}_k &\triangleq [r_k(0), r_k(T_s), \dots, r_k((L_o Q - 1)T_s)]^T \\ \mathbf{w}_k &\triangleq [w_k(0), w_k(T_s), \dots, w_k((L_o Q - 1)T_s)]^T \\ \mathbf{A}_{\tau_k} &\triangleq [\mathbf{a}_{-L_g}(\tau_k), \dots, \mathbf{a}_0(\tau_k), \dots, \mathbf{a}_{L_o+L_g-1}(\tau_k)] \\ \mathbf{a}_i(\tau_k) &\triangleq [g(-iT - \tau_k T), g(-iT + T_s - \tau_k T), \\ &\quad \dots, g(-iT + (L_o Q - 1)T_s - \tau_k T)]^T \\ \mathbf{s} &\triangleq [s(-L_g), \dots, s(0), \dots, s(L_o + L_g - 1)]^T. \end{aligned}$$

B. MISO Multiple Access Phase

Denote the data transmitted from the k th relay as $x_k(i)$, then the corresponding signal (within $0 \leq t \leq L_o T$) at the destination is

$$d_k(t) = h_k \sum_{i=-L_g}^{L_o+L_g-1} x_k(i)g(t-iT-\epsilon_k T). \quad (3)$$

The data are transmitted with $\mathbb{E}\{|x_k(i)|^2\} = E_x$ and $\epsilon_k \in [0, 1)$ is the unknown timing offset of the k th relay. At the destination, the observed waveform $d(t)$ is the superposition of the above waveforms from the K relays plus noise $d(t) = \sum_{k=1}^K d_k(t) + v(t)$, where $v(t)$ is the zero mean, circular complex Gaussian noise at the destination with variance σ_v^2 . After sampling with $Q \geq 2$ and putting the received samples into a vector $\mathbf{d} = [d(0), \dots, d((L_o Q - 1)T_s)]^T$, then we have

$$\mathbf{d} = \underbrace{[\mathbf{A}_{\epsilon_1} \mathbf{A}_{\epsilon_2} \cdots \mathbf{A}_{\epsilon_K}]}_{\triangleq \mathbf{A}_\epsilon} \underbrace{\begin{bmatrix} h_1 \mathbf{x}_1 \\ h_2 \mathbf{x}_2 \\ \vdots \\ h_K \mathbf{x}_K \end{bmatrix}}_{\triangleq \mathbf{X}} + \mathbf{v} = \mathbf{A}_\epsilon \mathbf{X} + \mathbf{v} \quad (4)$$

where $\mathbf{x}_k = [x_k(-L_g), \dots, x_k(L_o + L_g - 1)]^T$ is the data from the k th relay, and let $\epsilon \triangleq [\epsilon_1, \dots, \epsilon_K]^T$. The matrix \mathbf{A}_{ϵ_k} follows the same definition of \mathbf{A}_{τ_k} , and \mathbf{v} contains the samples of $v(t)$.

Remark: It is worth noting that the phase offsets between transmitters and receivers are not explicitly specified in the system model in (2) and (4), because they can be incorporated into the unknown channel coefficients while keeping the formulation of the system models unchanged.

C. Timing Synchronization Issue in Cooperative Systems

In the broadcasting phase, the synchronization problem at the relay is the same as the single user synchronization. Therefore, all the conventional synchronization techniques can be used [24], [27]. However, in the multiple access phase, asynchronous signals from different users overlap with each other, hence, there is no unique optimal sampling instant since the optimal instants for distinct users are different.

To cope with this problem, one way is to treat the signals from asynchronous users as signals passing through different paths in a multipath channel, then a joint maximum likelihood sequence estimator (JMLSE) can be used to jointly decode the data from all the users [11], [25]. Although JMLSE, in principle, can be applied to scenarios with any number of users, its implementation complexity becomes prohibitive when the number of users is greater than two.

A more direct method to resynchronize the multiple users is to design a resynchronization filter at the receiver to realign the asynchronous signals as first demonstrated in [26]. Unfortunately, general guidelines for choosing the optimal design parameters and comprehensive performance analysis studies are missing in [26]. In this paper, we will present a general framework for the resynchronization filter design and show that the scheme in [26] is a special case of the proposed framework. Since the timing and channel estimates are needed in the

resynchronization filter design, we first present the timing and channel estimation algorithms.

III. TIMING AND CHANNEL ESTIMATIONS

A. Maximum Likelihood Estimation at Relays

The joint timing and channel estimation in the first hop is performed at each relay individually, and it does not affect other relays' estimation performance. For the k th relay, the received signal is given by (2), and the likelihood function of the timing offset and the channel is

$$p(\mathbf{r}_k; \tau_k, c_k) = (\pi\sigma_{w_k}^2)^{-L_o Q} \exp\left\{-\frac{\|\mathbf{r}_k - c_k \mathbf{A}_{\tau_k} \mathbf{s}\|^2}{\sigma_{w_k}^2}\right\}$$

where \mathbf{s} is a known training sequence. Thus, maximizing the above function is equivalent to minimizing the following cost function:

$$J(\tau_k, c_k) = \|\mathbf{r}_k - c_k \mathbf{A}_{\tau_k} \mathbf{s}\|^2. \quad (5)$$

It can be readily shown that the ML estimate of the unknown channel coefficient c_k is (assuming τ_k is fixed)

$$\hat{c}_k = \frac{\mathbf{s}^H \mathbf{A}_{\tau_k}^H \mathbf{r}_k}{\mathbf{s}^H \mathbf{A}_{\tau_k}^H \mathbf{A}_{\tau_k} \mathbf{s}}. \quad (6)$$

After substituting (6) into (5) and ignoring some scaling constants and irrelevant terms, a cost function that only depends on τ_k is obtained as

$$\Lambda(\tau_k) = \|\mathbf{P}_\mathbf{A}^\perp(\tau_k) \mathbf{r}_k\|^2 \quad (7)$$

where $\mathbf{P}_\mathbf{A}^\perp(\tau_k) = \mathbf{I} - \mathbf{P}_\mathbf{A}(\tau_k)$ and $\mathbf{P}_\mathbf{A}(\tau_k) = \mathbf{A}_{\tau_k} \mathbf{s} \mathbf{s}^H \mathbf{A}_{\tau_k}^H / \mathbf{s}^H \mathbf{A}_{\tau_k}^H \mathbf{A}_{\tau_k} \mathbf{s}$. Then the timing offset τ_k is estimated as

$$\hat{\tau}_k = \arg \min_{\tau_k} \Lambda(\tau_k). \quad (8)$$

After the timing offset estimation, $\hat{\tau}_k$ is used to compensate the timing offset by using a matched filter with opposite delay $g(t + \hat{\tau}_k)$. Also, the channel coefficient can be readily obtained as $\hat{c}_k = \mathbf{s}^H \mathbf{A}_{\hat{\tau}_k}^H \mathbf{r}_k / \mathbf{s}^H \mathbf{A}_{\hat{\tau}_k}^H \mathbf{A}_{\hat{\tau}_k} \mathbf{s}$. This channel estimate is stored and used for decoding in the data transmission period.

B. Maximum Likelihood Estimation at Destination

In order to derive the joint timing and channel estimator in the multiple access phase, we start with an equivalent model of (4)

$$\mathbf{d} = \underbrace{[\mathbf{A}_{\epsilon_1} \mathbf{x}_1 \mathbf{A}_{\epsilon_2} \mathbf{x}_2 \cdots \mathbf{A}_{\epsilon_K} \mathbf{x}_K]}_{\triangleq \mathbf{\Omega}_\epsilon} \underbrace{\begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_K \end{bmatrix}}_{\triangleq \mathbf{h}} + \mathbf{v} \quad (9)$$

where \mathbf{x}_k denotes the distinct training sequence transmitted by the k th relay. In this study, we assume that the training sequences from different relays are uncorrelated (notice that these

sequences may not be optimal, but the problem of training design is beyond the scope of this paper).

The joint likelihood function of the timing offsets and channels is given by

$$p(\mathbf{d}; \boldsymbol{\epsilon}, \mathbf{h}) = (\pi\sigma_v^2)^{-L_o Q} \exp \left\{ -\frac{\|\mathbf{d} - \boldsymbol{\Omega}_\epsilon \mathbf{h}\|^2}{\sigma_v^2} \right\}.$$

Similarly, maximizing the above function is equivalent to minimizing the following cost function:

$$J(\boldsymbol{\epsilon}, \mathbf{h}) = \|\mathbf{d} - \boldsymbol{\Omega}_\epsilon \mathbf{h}\|^2. \quad (10)$$

Moreover, by substituting $\hat{\mathbf{h}} = (\boldsymbol{\Omega}_\epsilon^H \boldsymbol{\Omega}_\epsilon)^{-1} \boldsymbol{\Omega}_\epsilon^H \mathbf{d}$ back into the cost function and dropping some constant terms, we obtain the compressed log-likelihood function

$$\Lambda(\boldsymbol{\epsilon}) = \|\mathbf{P}_\Omega(\boldsymbol{\epsilon}) \mathbf{d}\|^2 \quad (11)$$

where $\mathbf{P}_\Omega(\boldsymbol{\epsilon}) = \boldsymbol{\Omega}_\epsilon (\boldsymbol{\Omega}_\epsilon^H \boldsymbol{\Omega}_\epsilon)^{-1} \boldsymbol{\Omega}_\epsilon^H$. Then the multiple timing offsets $\boldsymbol{\epsilon}$ are estimated as

$$\hat{\boldsymbol{\epsilon}} = \arg \max_{\boldsymbol{\epsilon}} \Lambda(\boldsymbol{\epsilon}). \quad (12)$$

Finally the channels are estimated using the estimates of $\boldsymbol{\epsilon}$

$$\hat{\mathbf{h}} = (\boldsymbol{\Omega}_{\hat{\boldsymbol{\epsilon}}}^H \boldsymbol{\Omega}_{\hat{\boldsymbol{\epsilon}}})^{-1} \boldsymbol{\Omega}_{\hat{\boldsymbol{\epsilon}}} \mathbf{d}. \quad (13)$$

In the above maximization problem (12), there are multiple timing offsets from different relay-destination pairs, which would require an exhaustive search over a multidimensional space, imposing great computational complexity at the receiver. To overcome this problem, *alternating projection* can be used to reduce the K -dimensional maximization into a series of 1-D maximization problems. Details on the implementation of alternating projection is omitted here since it has been elaborated in [28].

IV. CRB ANALYSIS

Traditionally, CRBs are used as performance benchmarks for any unbiased estimators. In synchronization where channel estimation does not play a dominant role, Conditional CRB (CCRB) [27] is usually used to evaluate the performance of the timing estimator. However, strictly speaking, CRB or CCRB are not valid bounds at low SNR for synchronization parameters because the derivation of CRB and CCRB do not assume any prior information on the parameter, while in fact, the range of the timing offset $[0, 1)$ is informative in that the largest mean-square error (MSE) possibly achieved by any timing estimator is on the order of 10^{-1} . On the contrary, notice that channel estimation errors are well bounded by the CRB because there is no such range limit for any channel realization.

The reason why the CRB or CCRB can serve as the performance bounds for synchronization problems in previous studies [24], [27] is that the CRBs/CCRBs are derived for single user scenario and generally nonfading environments, in which the limitation of CRBs has not yet begun to show at practical SNR (medium to high SNR) levels. However, for synchronization in fading channels, the equivalent SNR under deep fading is

very low, hence, the CRB becomes inapplicable in this situation. Here we develop a modification on the CRB, called the Weighted Bayesian CRB (WBCRB), for the timing estimation to incorporate the prior information on the range of the timing offset. WBCRB is more complicated than CRB but it is a valid bound even in fading situations at any SNR. Since the derivation of CRB represents the theoretical foundation for the derivation of WBCRB, and the CRB for channels is still a valid bound, we first derive the joint CRB for timing and channel estimation without assuming any prior information for the timing parameters.

A. Joint CRBs

Since the CRB for the timing and channel estimation at the relay during SIMO broadcasting phase is the same as that in the SISO case, it can be viewed as a special case of the MISO multiple access phase when the number of relay equals to one. Hence the joint CRB is derived for the MISO phase and the result for SIMO phase is evaluated as a special case. Recall the signal model in the second phase transmission given in (9) as $\mathbf{d} = \boldsymbol{\Omega}_\epsilon \mathbf{h} + \mathbf{v}$. Let $\boldsymbol{\theta} \triangleq [\boldsymbol{\epsilon}^T, \Re\{\mathbf{h}\}^T, \Im\{\mathbf{h}\}^T]^T$ and denote $\boldsymbol{\mu} = \boldsymbol{\Omega}_\epsilon \mathbf{h}$. The Fisher Information Matrix (FIM) \mathbf{F} of $\boldsymbol{\theta}$ is given by [29]

$$\mathbf{F} = \frac{1}{\sigma_v^2} \cdot \Re \left\{ \frac{\partial \boldsymbol{\mu}^H}{\partial \boldsymbol{\theta}} \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}^T} \right\}$$

which is then readily obtained as

$$\mathbf{F} = \frac{1}{\sigma_v^2} \times \begin{bmatrix} \Re \{ \mathbf{H}^H \boldsymbol{\Psi}_\epsilon^H \boldsymbol{\Psi}_\epsilon \mathbf{H} \} & \Re \{ \mathbf{H}^H \boldsymbol{\Psi}_\epsilon^H \boldsymbol{\Omega}_\epsilon \} & \Im \{ \mathbf{H}^H \boldsymbol{\Psi}_\epsilon^H \boldsymbol{\Omega}_\epsilon \} \\ \Re \{ \boldsymbol{\Omega}_\epsilon^H \boldsymbol{\Psi}_\epsilon \mathbf{H} \} & \Re \{ \boldsymbol{\Omega}_\epsilon^H \boldsymbol{\Omega}_\epsilon \} & \Im \{ \boldsymbol{\Omega}_\epsilon^H \boldsymbol{\Omega}_\epsilon \} \\ -\Im \{ \boldsymbol{\Omega}_\epsilon^H \boldsymbol{\Psi}_\epsilon \mathbf{H} \} & -\Im \{ \boldsymbol{\Omega}_\epsilon^H \boldsymbol{\Omega}_\epsilon \} & \Re \{ \boldsymbol{\Omega}_\epsilon^H \boldsymbol{\Omega}_\epsilon \} \end{bmatrix}$$

where $\boldsymbol{\Psi}_\epsilon = [\mathbf{D}_{\epsilon_1} \mathbf{x}_1, \mathbf{D}_{\epsilon_2} \mathbf{x}_2, \dots, \mathbf{D}_{\epsilon_K} \mathbf{x}_K]$ with $\mathbf{D}_{\epsilon_k} = \partial \mathbf{A}_{\epsilon_k} / \partial \epsilon_k$, and $\mathbf{H} = \text{diag}(\mathbf{h})$. The CRB is finally computed through similar mathematical derivations as in [29], and assumes the expressions

$$\text{CRB}_{\text{MA}}(\boldsymbol{\epsilon}) = \sigma_v^2 \cdot (\Re \{ \mathbf{H}^H \boldsymbol{\Psi}_\epsilon^H \mathbf{P}_\Omega^\perp(\boldsymbol{\epsilon}) \boldsymbol{\Psi}_\epsilon \mathbf{H} \})^{-1} \quad (14)$$

$$\begin{aligned} \text{CRB}_{\text{MA}}(\mathbf{h}) &= \sigma_v^2 \cdot \left(2(\boldsymbol{\Omega}_\epsilon^H \boldsymbol{\Omega}_\epsilon)^{-1} + (\boldsymbol{\Omega}_\epsilon^H \boldsymbol{\Omega}_\epsilon)^{-1} \boldsymbol{\Omega}_\epsilon^H \boldsymbol{\Psi}_\epsilon \mathbf{H} \right. \\ &\quad \times (\Re \{ \mathbf{H}^H \boldsymbol{\Psi}_\epsilon^H \mathbf{P}_\Omega^\perp(\boldsymbol{\epsilon}) \boldsymbol{\Psi}_\epsilon \mathbf{H} \})^{-1} \\ &\quad \left. \times \mathbf{H}^H \boldsymbol{\Psi}_\epsilon^H \boldsymbol{\Omega}_\epsilon (\boldsymbol{\Omega}_\epsilon^H \boldsymbol{\Omega}_\epsilon)^{-1} \right) \end{aligned} \quad (15)$$

where the subscript ‘‘MA’’ stands for the multiple access phase. By setting the number of relays to one and replacing the expressions with their counterparts in the SIMO broadcasting phase, the joint CRB for the k th relay in the SIMO broadcasting phase is

$$\text{CRB}_{\text{BC}}(\tau_k) = \frac{\sigma_{w_k}^2}{|c_k|^2 \mathbf{s}^H \mathbf{D}_{\tau_k}^H \mathbf{P}_\mathbf{A}^\perp(\tau_k) \mathbf{D}_{\tau_k} \mathbf{s}} \quad (16)$$

$$\begin{aligned} \text{CRB}_{\text{BC}}(c_k) &= \frac{2\sigma_{w_k}^2}{\mathbf{s}^H \mathbf{A}_{\tau_k}^H \mathbf{A}_{\tau_k} \mathbf{s}} + |c_k|^2 \frac{\mathbf{s}^H \mathbf{A}_{\tau_k}^H \mathbf{D}_{\tau_k} \mathbf{s}}{\mathbf{s}^H \mathbf{A}_{\tau_k}^H \mathbf{A}_{\tau_k} \mathbf{s}} \\ &\quad \times \text{CRB}_{\text{BC}}(\tau_k) \cdot \left(\frac{\mathbf{s}^H \mathbf{A}_{\tau_k}^H \mathbf{D}_{\tau_k} \mathbf{s}}{\mathbf{s}^H \mathbf{A}_{\tau_k}^H \mathbf{A}_{\tau_k} \mathbf{s}} \right)^H \end{aligned} \quad (17)$$

where the matrix $\mathbf{P}_{\mathbf{A}}^{\perp}(\tau_k)$ is defined in (7).

B. Weighted Bayesian CRB

As discussed at the beginning of this section, the CRB derived above for the timing estimate is not applicable in the low SNR region due to the fact that the given finite range on the timing estimates serves as an informative prior during estimation [30].

The most common bound that considers the prior information of the parameter of interest is the Bayesian CRB (BCRB). However as demonstrated in [30], BCRB does not exist for parameters with uniform distribution. In order to derive a valid performance bound for parameters with a given finite range and uniform distribution, the WBCRB was introduced and studied in [30]. The WBCRB can be shown to be a valid lower bound and evaluated in [31] as

$$\begin{aligned} \text{WBCRB}_{\text{MA}}(\epsilon) &= \mathbb{E}_{\epsilon} \{ \mathbf{Q}(\epsilon) \} \{ \mathbb{E}_{\epsilon} \{ \mathbf{F}_w(\epsilon) \} + \mathbb{E}_{\epsilon} \{ \mathbf{P}_w(\epsilon) \} \}^{-1} \mathbb{E}_{\epsilon} \{ \mathbf{Q}(\epsilon) \} \end{aligned}$$

where $\mathbf{F}_w(\epsilon)$ is the weighted FIM for the timing offset parameters and $\mathbf{P}_w(\epsilon)$ is the weighted Prior Information Matrix (PIM), which are defined in (21) and (20), respectively. The symbol $\mathbf{Q}(\epsilon) \triangleq \text{diag}(q(\epsilon_1), \dots, q(\epsilon_K))$ represents the weighting matrix, with $q(\epsilon_k)$ being the individual weighting function for the timing offset of the k th relay. As suggested by [30], the weighting function $q(\epsilon_k)$ is chosen as

$$q(\epsilon_k) = \begin{cases} \epsilon_k^{\gamma} (1 - \epsilon_k)^{\gamma}, & 0 \leq \epsilon_k < 1 \\ 0, & \text{otherwise} \end{cases} \quad (18)$$

where γ is the *weighting index*. The value of γ is chosen to adjust the tightness of the WBCRB, and the optimal value of γ can only be determined numerically [30].

1) *Calculation of $\mathbb{E}_{\epsilon} \{ \mathbf{Q}(\epsilon) \}$ and $\mathbb{E}_{\epsilon} \{ \mathbf{P}_w(\epsilon) \}$* : The evaluation of $\mathbb{E}_{\epsilon} \{ \mathbf{Q}(\epsilon) \}$ can be obtained easily as an extension of the derivation in [30], and leads to the following result:

$$\mathbb{E}_{\epsilon} \{ \mathbf{Q}(\epsilon) \} = -\beta(\gamma + 1, \gamma + 1) \mathbf{I} \quad (19)$$

where $\beta(\cdot, \cdot)$ denotes the beta function $\beta(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx$. Meanwhile, the weighted PIM $\mathbf{P}_w(\epsilon)$ is defined in [31] as

$$\begin{aligned} [\mathbf{P}_w(\epsilon)]_{i,j} &= \left(q(\epsilon_i) q(\epsilon_j) \frac{\partial \ln[q(\epsilon_i) p(\epsilon)]}{\partial \epsilon_i} \cdot \frac{\partial \ln[q(\epsilon_j) p(\epsilon)]}{\partial \epsilon_j} \right) \end{aligned} \quad (20)$$

where $p(\epsilon)$ is the prior distribution of timing offsets. With ϵ_k being uniform distributed in $[0, 1)$ and the weighting function $q(\epsilon_k)$ defined in (18), it can be shown that [30]

$$\mathbb{E}_{\epsilon} \{ [\mathbf{P}_w(\epsilon)]_{i,j} \} = \begin{cases} \gamma \cdot \beta(2\gamma - 1, 2\gamma + 1), & i = j \\ 0, & i \neq j. \end{cases}$$

2) *Calculation of $\mathbf{F}_w(\epsilon)$* : The weighted FIM for the timing offset parameters is defined in [31] as

$$[\mathbf{F}_w(\epsilon)]_{i,j} = \left(q(\epsilon_i) q(\epsilon_j) \frac{\partial \ln p(\mathbf{d}|\epsilon)}{\partial \epsilon_i} \cdot \frac{\partial \ln p(\mathbf{d}|\epsilon)}{\partial \epsilon_j} \right) \quad (21)$$

where the vector \mathbf{d} represents the received samples in (9). Notice that $p(\mathbf{d}|\epsilon)$ is the conditional probability distribution of \mathbf{d} given ϵ . However, in the joint timing and channel estimation problem considered herein, \mathbf{h} is modeled as deterministic unknown, so we only have $p(\mathbf{d}|\epsilon, \mathbf{h})$. In order to eliminate the nuisance parameter \mathbf{h} , we employ the method introduced in [27] of using a conditional approach to asymptotically (i.e., K is large) obtain the distribution by substituting the estimate of $\hat{\mathbf{h}}$ in (13) back into the joint distribution function $p(\mathbf{d}|\epsilon, \mathbf{h})$. Following this logic, it can be readily shown [27] that

$$\frac{\partial \ln p(\mathbf{d}|\epsilon)}{\partial \epsilon_i} \cdot \frac{\partial \ln p(\mathbf{d}|\epsilon)}{\partial \epsilon_j} \simeq \frac{1}{\sigma_v^2} \left[\Re \{ \mathbf{H}^H \Psi_{\epsilon}^H \mathbf{P}_{\Omega}^{\perp}(\epsilon) \Psi_{\epsilon} \mathbf{H} \} \right]_{i,j}.$$

Last but not least, even though the expression $\mathbf{F}_w(\epsilon)$ is in closed form, the value of $\mathbb{E}_{\epsilon} \{ \mathbf{F}_w(\epsilon) \}$ can only be obtained numerically because it depends on the pulse shaping filter $g(t)$ in a very complicated and mathematically intractable manner. It is also worth noting that if $\gamma = 0$, the WBCRB has the following relationship with CRB in (14)

$$\text{WBCRB}_{\text{MA}}(\epsilon)|_{\gamma=0} = \left(\mathbb{E}_{\epsilon} \{ \text{CRB}_{\text{MA}}^{-1}(\epsilon) \} \right)^{-1}. \quad (22)$$

By setting the number of relays to one and replacing the expressions with their counterparts in the SIMO broadcasting phase, the WBCRB for the k th relay in the SIMO broadcasting phase is

$$\begin{aligned} \text{WBCRB}_{\text{BC}}(\tau_k) &= \frac{\beta^2(\gamma + 1, \gamma + 1)}{\mathbb{E}_{\tau_k} \{ F_w(\tau_k) \} / \sigma_{w_k}^2 + \gamma \cdot \beta(2\gamma - 1, 2\gamma + 1)} \end{aligned}$$

where $F_w(\tau_k) = q^2(\tau_k) |c_k|^2 \mathbf{s}^H \mathbf{D}_{\tau_k}^H \mathbf{P}_{\mathbf{A}}^{\perp}(\tau_k) \mathbf{D}_{\tau_k} \mathbf{s}$.

Finally, note that the WBCRB derived in this paper is novel and different from the results available in [27], [30], and [31] because there exist multiple nuisance parameters in the current signal model (the channel coefficients), and also the bound is with respect to multiple timing parameters. Comparisons between CRB and WBCRB will be provided in Section VI (Figs. 6 and 7).

V. RESYNCHRONIZATION FILTER DESIGN FOR THE MULTIPLE ACCESS PHASE

At the destination node, after the multiple timing offsets ϵ have been estimated through the proposed estimator, the remaining issue is how to employ the estimates to compensate the offsets to obtain ISI-free reception. Mathematically, this problem can be solved by designing a resynchronization filter \mathbf{f} such that after the convolution with the received signal (4), the ISI components are eliminated. This process is actually

an equalization for the fractional timing offset distortion and equivalent to the minimization problem

$$\min_{\mathbf{f}} \mathbb{E} \left\{ \left\| \hat{\mathbf{X}}^H (\mathbf{A}_{\hat{\epsilon}}^H \mathbf{f} - \mathbf{b}) \right\|^2 \right\} \quad (23)$$

where the expectation is taken with respect to data realization. The vector $\mathbf{b} \triangleq [\delta_1^H, \dots, \delta_K^H]^H$ and $\delta_1 = \dots = \delta_K \triangleq [R_{gg}(-L_g - M_o), \dots, R_{gg}(0), \dots, R_{gg}(L_g + M_o)]^H$ stands for the ideal zero-ISI sampled waveform after matched filtering, where $M_o = (L_o - 1)/2$ (for the sake of discussion, L_o is taken to be an odd number) and $R_{gg}(\tau)$ is the autocorrelation function of $g(t)$ at $t = \tau T$. The expression of $\hat{\mathbf{X}}$ in (23) is the \mathbf{X} in (4) with h_i replaced by \hat{h}_i and can be equivalently expressed as

$$\hat{\mathbf{X}} = \underbrace{\begin{bmatrix} \hat{h}_1 \mathbf{I} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \hat{h}_K \mathbf{I} \end{bmatrix}}_{\triangleq \mathcal{H}} \underbrace{\begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_K \end{bmatrix}}_{\triangleq \mathcal{X}} = \mathcal{H}\mathcal{X}. \quad (24)$$

Thus, (23) can be reformulated as the following weighted least squares (WLS) problem

$$\min_{\mathbf{f}} (\mathbf{A}_{\hat{\epsilon}}^H \mathbf{f} - \mathbf{b})^H \mathbf{\Pi} (\mathbf{A}_{\hat{\epsilon}}^H \mathbf{f} - \mathbf{b}). \quad (25)$$

The symbol $\mathbf{\Pi} \triangleq \mathcal{H} \mathbf{R}_{\mathbf{x}} \mathcal{H}^H$ represents the positive definite $m \times m$ weighting matrix with $m = K(L_o + 2L_g)$, and $\mathbf{R}_{\mathbf{x}} = \mathbb{E}\{\mathcal{X}\mathcal{X}^H\}$ is given by

$$\mathbf{R}_{\mathbf{x}} \triangleq \begin{bmatrix} \mathbf{R}_{1,1} & \cdots & \mathbf{R}_{1,K} \\ \vdots & \ddots & \vdots \\ \mathbf{R}_{K,1} & \cdots & \mathbf{R}_{K,K} \end{bmatrix}$$

with the submatrix $\mathbf{R}_{i,j} = \mathbb{E}\{\mathbf{x}_i \mathbf{x}_j^H\}$ being the correlation matrix between \mathbf{x}_i and \mathbf{x}_j . Note that once the transmission scheme and cooperative strategy are fixed, the correlation matrix can be determined and made available at both the transmitter and the receiver.

Generally speaking, the solution to (25) is the WLS solution $\mathbf{f} = (\mathbf{A}_{\hat{\epsilon}}^H \mathbf{\Pi} \mathbf{A}_{\hat{\epsilon}}^H)^{-1} \mathbf{A}_{\hat{\epsilon}}^H \mathbf{\Pi} \mathbf{b}$. However, the timing offsets in the MISO phase can be represented as $\epsilon_k = \epsilon_o + \Delta_k$, where ϵ_o is the common offset with respect to a certain time frame at the destination and Δ_k is the residual offset between the individual offset and the common offset, including the travel delay, processing latency and the timing estimation error in the SIMO broadcasting phase at each relay. Therefore, the columns of $\mathbf{A}_{\hat{\epsilon}}^H$ are quite similar one with respect to the other, and hence, (25) is an ill-posed problem. Unlike a rank deficient problem which can be solved by discarding the zero or close-to-zero singular values and corresponding singular vectors, for an ill-conditioned matrix there are no general rules to determine which singular values to discard [32] as there is no significant gap between the singular values (an example of the singular value distribution of $\mathbf{A}_{\hat{\epsilon}}^H$ is shown in Fig. 3), leading to the amplification of any perturbation that exists in the system.

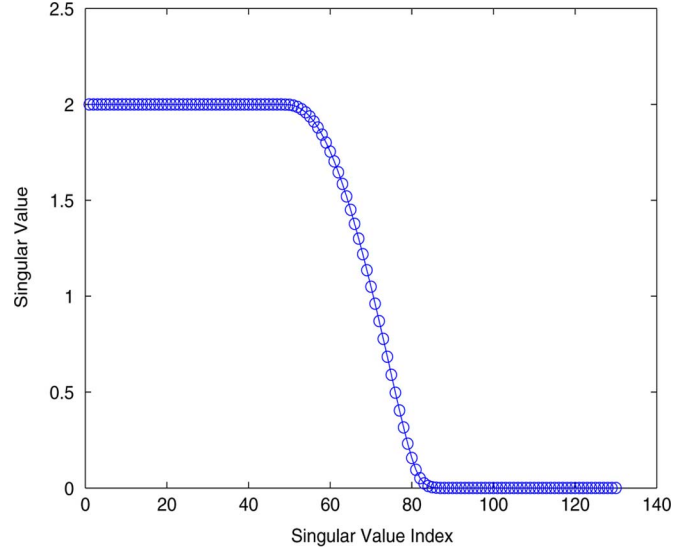


Fig. 3. Example of the singular values distribution of the ill-conditioned matrix $\mathbf{A}_{\hat{\epsilon}}$ with $L_o = 65$, $L_g = 4$ and $Q = 2$.

In the following, we present two methods to deal with the filter design problem mentioned above. One is to bypass the ill-posed problem by employing a transmitter pulse shaping filter to compensate a nominal delay that minimizes the residual ISI power, while the other is a more general framework that employs the regularization theory to design a regularized filter with arbitrary shapes.

A. Matched Filtering

To bypass the ill-conditioned formulation, a natural choice of the resynchronization filter \mathbf{f} would be to use a transmitter pulse shaping filter $g(t + \epsilon_c)$ with a nominal delay ϵ_c . If there is only a single timing offset, the mismatch can be perfectly compensated by a matched filter with a nominal delay ϵ_c that is opposite to the offset. When there are multiple timing offsets, the nominal delay is chosen to minimize the residual ISI power at the output. The residual ISI after resynchronization filtering can be expressed as follows:

$$P_{\text{ISI}} = (\mathbf{A}_{\hat{\epsilon}}^H \mathbf{f} - \mathbf{b})^H \mathbf{\Pi} (\mathbf{A}_{\hat{\epsilon}}^H \mathbf{f} - \mathbf{b}). \quad (26)$$

The quantity $(\mathbf{A}_{\hat{\epsilon}}^H \mathbf{f} - \mathbf{b})$ is a vector including the residual ISI errors from different users. Let $\mathbf{A}_{\hat{\epsilon}}^H \mathbf{f} - \mathbf{b} \triangleq [\mathbf{e}_1^H, \dots, \mathbf{e}_K^H]^H$ for notation simplicity. Now P_{ISI} in (26) can be expanded as

$$P_{\text{ISI}} = \sum_{k_1=1}^K \sum_{k_2=1}^K \hat{h}_{k_1} \hat{h}_{k_2}^* \mathbf{e}_{k_1}^H \mathbf{R}_{k_1, k_2} \mathbf{e}_{k_2}. \quad (27)$$

The residual ISI error vector \mathbf{e}_k from the k th user can be easily calculated as

$$\mathbf{e}_k = \begin{bmatrix} R_{gg}(-L_g - M_o + \epsilon_c - \hat{\epsilon}_k) - R_{gg}(-L_g - M_o) \\ \vdots \\ R_{gg}(\epsilon_c - \hat{\epsilon}_k) - R_{gg}(0) \\ \vdots \\ R_{gg}(L_g + M_o + \epsilon_c - \hat{\epsilon}_k) - R_{gg}(L_g + M_o) \end{bmatrix}.$$

Since $\epsilon_c - \hat{\epsilon}_k$ is very small, by applying Taylor series expansion on \mathbf{e}_k , we have

$$\mathbf{e}_k = \begin{bmatrix} R'_{gg}(-L_g - M_o)(\epsilon_c - \hat{\epsilon}_k) \\ + \frac{1}{2} R''_{gg}(-L_g - M_o)(\epsilon_c - \hat{\epsilon}_k)^2 + O((\epsilon_c - \hat{\epsilon}_k)^3) \\ \vdots \\ R'_{gg}(0)(\epsilon_c - \hat{\epsilon}_k) \\ + \frac{1}{2} R''_{gg}(0)(\epsilon_c - \hat{\epsilon}_k)^2 + O((\epsilon_c - \hat{\epsilon}_k)^3) \\ \vdots \\ R'_{gg}(L_g + M_o)(\epsilon_c - \hat{\epsilon}_k) \\ + \frac{1}{2} R''_{gg}(L_g + M_o)(\epsilon_c - \hat{\epsilon}_k)^2 + O((\epsilon_c - \hat{\epsilon}_k)^3) \end{bmatrix}$$

where $O((\epsilon_c - \hat{\epsilon}_k)^3)$ represents the higher-order terms. Then by expanding $\mathbf{e}_{k_1}^H \mathbf{R}_{k_1, k_2} \mathbf{e}_{k_2}$ and omitting the terms with order higher than $O((\epsilon_c - \hat{\epsilon}_k)^3)$, we obtain

$$\mathbf{e}_{k_1}^H \mathbf{R}_{k_1, k_2} \mathbf{e}_{k_2} \cong (\epsilon_c - \hat{\epsilon}_{k_1})(\epsilon_c - \hat{\epsilon}_{k_2})f(k_1, k_2) \quad (28)$$

where

$$f(k_1, k_2) \triangleq \sum_{i=-L_g-M_o}^{L_g+M_o} \sum_{j=-L_g-M_o}^{L_g+M_o} \mathbf{R}_{k_1, k_2}^{(i, j)} R'_{gg}(i) R'_{gg}(j)$$

and $\mathbf{R}_{k_1, k_2}^{(i, j)}$ represents the (i, j) th element in the correlation matrix \mathbf{R}_{k_1, k_2} . Hence (27) becomes

$$P_{\text{ISI}}(\epsilon_c) = \sum_{k_1=1}^K \sum_{k_2=1}^K \hat{h}_{k_1} \hat{h}_{k_2}^* f(k_1, k_2) \times [\epsilon_c^2 - \epsilon_c(\hat{\epsilon}_{k_1} + \hat{\epsilon}_{k_2}) + \hat{\epsilon}_{k_1} \hat{\epsilon}_{k_2}]$$

which is a quadratic function of ϵ_c . After rearranging the terms, we have

$$P_{\text{ISI}} = A_{\Sigma} \cdot \epsilon_c^2 - B_{\Sigma} \cdot \epsilon_c + C_{\Sigma} \quad (29)$$

where

$$\begin{aligned} A_{\Sigma} &= \sum_{k_1=1}^K \sum_{k_2=1}^K \hat{h}_{k_1} \hat{h}_{k_2}^* f(k_1, k_2) \\ B_{\Sigma} &= \sum_{k_1=1}^K \sum_{k_2=1}^K \hat{h}_{k_1} \hat{h}_{k_2}^* f(k_1, k_2) (\hat{\epsilon}_{k_1} + \hat{\epsilon}_{k_2}) \\ C_{\Sigma} &= \sum_{k_1=1}^K \sum_{k_2=1}^K \hat{h}_{k_1} \hat{h}_{k_2}^* f(k_1, k_2) \hat{\epsilon}_{k_1} \hat{\epsilon}_{k_2}. \end{aligned}$$

It is shown in the Appendix that $A_{\Sigma} > 0$, therefore the minimum of (29) is achieved when $\epsilon_c = B_{\Sigma}/2A_{\Sigma}$. In particular, when $\mathbf{R}_{\mathbf{x}} \propto \mathbf{I}$, we have

$$\epsilon_c = \frac{\sum_{k=1}^K |\hat{h}_k|^2 \hat{\epsilon}_k}{\sum_{k=1}^K |\hat{h}_k|^2} \quad (30)$$

which means that a matched filter that shifts back the weighted mean delay of the signals is optimal.

B. Regularized Filtering

The method of employing the transmitter pulse shaping filter is intuitively appealing, however it presents limited performance

due to the filter shape constraint. Therefore, we hereby propose the *weighted regularized least squares* (WRLS) [33] (also known as *scaled Tikhonov regularization*) method to deal with the ill-posed problem in (25). To simplify our notation in the following analysis, the optimization problem in (25) is rewritten compactly as

$$\min_{\mathbf{f}} (\mathbf{G}_{\Pi} \mathbf{f} - \mathbf{b}_{\Pi})^H (\mathbf{G}_{\Pi} \mathbf{f} - \mathbf{b}_{\Pi}) \quad (31)$$

where $\mathbf{G}_{\Pi} = \mathbf{\Pi}^{\frac{1}{2}} \mathbf{A}_{\epsilon}^H$ and $\mathbf{b}_{\Pi} = \mathbf{\Pi}^{\frac{1}{2}} \mathbf{b}$, with $\mathbf{\Pi}^{\frac{1}{2}}$ representing the Cholesky square root of the weighting matrix $\mathbf{\Pi}$. Now the WRLS problem can be stated as

$$\min_{\mathbf{f}} [(\mathbf{G}_{\Pi} \mathbf{f} - \mathbf{b}_{\Pi})^H (\mathbf{G}_{\Pi} \mathbf{f} - \mathbf{b}_{\Pi}) + \lambda^2 \mathbf{f}^H \mathbf{L}^H \mathbf{L} \mathbf{f}] \quad (32)$$

where $\mathbf{f}^H \mathbf{L}^H \mathbf{L} \mathbf{f}$ controls the properties of the regularized solution norm by choosing different regularization matrices \mathbf{L} (e.g., identity matrix for minimum energy, first derivative for maximum flatness [32]); the variable λ stands for the regularization parameter that balances the minimization of the two terms. When $\lambda = 0$, formulation (32) reduces to the WLS problem in (25). In order to find the solution of (32), we rewrite (32) as

$$\min_{\mathbf{f}} \left\| \begin{pmatrix} \mathbf{G}_{\Pi} \\ \lambda \mathbf{L} \end{pmatrix} \mathbf{f} - \begin{pmatrix} \mathbf{b}_{\Pi} \\ \mathbf{0} \end{pmatrix} \right\|^2.$$

For a fixed λ , the solution is readily obtained as

$$\mathbf{f}_{\lambda} = (\mathbf{G}_{\Pi}^H \mathbf{G}_{\Pi} + \lambda^2 \mathbf{L}^H \mathbf{L})^{-1} \mathbf{G}_{\Pi}^H \mathbf{b}_{\Pi}. \quad (33)$$

Notice that in (33), if λ is too large, the solution is overregularized and the residual error may be overwhelming, while if it is too small, the solution becomes underregularized and the stability of the solution is highly affected. As can be seen, the performance of the regularized solution \mathbf{f}_{λ} depends heavily on the *regularization parameter* λ (the choice of \mathbf{L} is relatively straightforward as different \mathbf{L} present different physical meanings).

1) *Choice of Regularization Parameter λ* : As discussed previously, the value of λ could significantly affect the performance of the resynchronization filter. Therefore, choosing an appropriate λ is very important. A conceptually simple tool for the analysis of discrete ill-posed problems is the *L-curve* technique [32], which is a plot of the discrete smoothing norm $\|\mathbf{L} \mathbf{f}_{\lambda}\|^2$ of the regularized solution versus the corresponding residual norm $\|\mathbf{G}_{\Pi} \mathbf{f}_{\lambda} - \mathbf{b}_{\Pi}\|^2$ for different λ . The L-curve clearly displays the compromise between minimization of these two quantities as illustrated in Fig. 4 (with $\mathbf{L} = \mathbf{I}$).

The L-curve corner, defined as the *maximum curvature* point on the curve $(\log \|\mathbf{G}_{\Pi} \mathbf{f}_{\lambda} - \mathbf{b}_{\Pi}\|, \log \|\mathbf{L} \mathbf{f}_{\lambda}\|)$, appears to be a good compromise that balances the regularization errors and perturbation errors in the regularized solution \mathbf{f}_{λ} [32]. Therefore, we will take this point for the regularization parameter λ . In order to find the λ that corresponds to the L-curve corner, let $\rho = \|\mathbf{G}_{\Pi} \mathbf{f}_{\lambda} - \mathbf{b}_{\Pi}\|$ and $\xi = \|\mathbf{L} \mathbf{f}_{\lambda}\|$. Then the curvature $\kappa(\lambda)$ is given by [32]

$$\kappa(\lambda) = 2 \frac{\xi \rho}{\xi'} \frac{\lambda^2 \xi' \rho + 2 \lambda \xi \rho + \lambda^4 \xi \xi'}{(\lambda^2 \xi'^2 + \rho^2)^{3/2}} \quad (34)$$

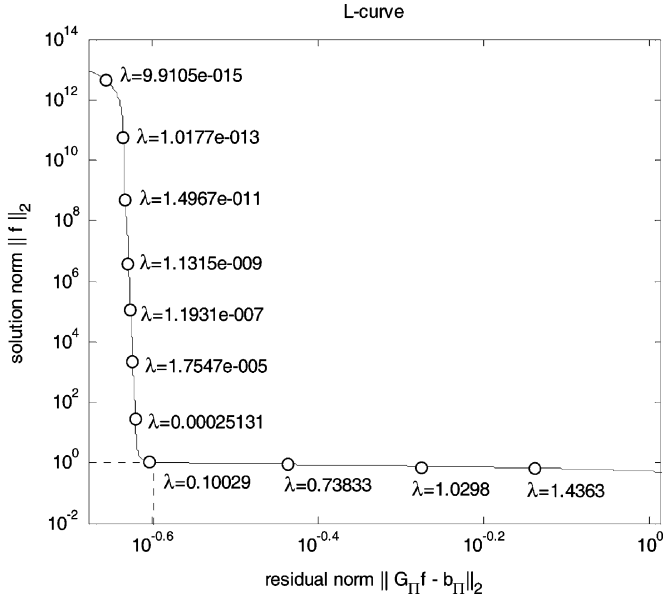


Fig. 4. An example of L-curve.

where

$$\xi' = \frac{4}{\lambda} \mathbf{f}_\lambda^H (\mathbf{G}_\Pi^H \mathbf{G}_\Pi + \lambda^2 \mathbf{L}^H \mathbf{L})^{-1} \mathbf{G}_\Pi^H (\mathbf{G}_\Pi \mathbf{f}_\lambda - \mathbf{b}_\Pi). \quad (35)$$

Note that when channel and timing offset realizations are given, ρ, ξ and ξ' can be computed. Therefore, the regularization parameter λ with maximum curvature in (34) can be located (e.g., by exhaustive search).

2) *Special Case—Least Squares With a Quadratic Constraint*: The resynchronization filter design was first discussed in [26] using the LS with a quadratic constraint (the Fixed SNR Loss or the Fixed Fidelity Loss methods). Note that there is no weighting matrix $\mathbf{\Pi}$ in the original method in [26]. In the following, we incorporate the weighting matrix $\mathbf{\Pi}$ to make the discussion more general, and setting $\mathbf{\Pi} = \mathbf{I}$ will reduce the formulation back to that of [26]. The methods in [26] can be stated as

- Fixed SNR Loss Requirement

$$\min \|\mathbf{G}_\Pi \mathbf{f} - \mathbf{b}_\Pi\|^2 \quad \text{s.t.} \quad \|\mathbf{L} \mathbf{f}\|^2 \leq \alpha \quad (36)$$

- Fixed Fidelity Loss Requirement

$$\min \|\mathbf{L} \mathbf{f}\|^2 \quad \text{s.t.} \quad \|\mathbf{G}_\Pi \mathbf{f} - \mathbf{b}_\Pi\|^2 \leq \beta. \quad (37)$$

The constraints α and β are design parameters that bound the norm of the solution and the errors, respectively. In the following, we derive the relationship between the regularization framework and these two methods. In order to do this, the generalized singular value decomposition (GSVD) of the matrix pair $(\mathbf{G}_\Pi, \mathbf{L})$ is needed

$$\mathbf{G}_\Pi = \mathbf{U} \mathbf{\Sigma} \mathbf{Y}^{-1}, \quad \mathbf{L} = \mathbf{V} \mathbf{M} \mathbf{Y}^{-1} \quad (38)$$

where $\mathbf{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_n)$, $\mathbf{M} = \text{diag}(\mu_1, \dots, \mu_n)$ are the singular values of \mathbf{G}_Π and \mathbf{L} , respectively, with $n = L_o Q$. Matrices \mathbf{U} and \mathbf{V} represent the unitary singular vector matrices of \mathbf{G}_Π and \mathbf{L} (i.e., $\mathbf{U}^H \mathbf{U} = \mathbf{V}^H \mathbf{V} = \mathbf{I}$), while \mathbf{Y} is

a nonsingular matrix that satisfies $\mathbf{Y}^H \mathbf{G}_\Pi^H \mathbf{G}_\Pi \mathbf{Y} = \mathbf{\Sigma}^2$ and $\mathbf{Y}^H \mathbf{L}^H \mathbf{L} \mathbf{Y} = \mathbf{M}^2$. With the GSVD of $(\mathbf{G}_\Pi, \mathbf{L})$, the regularized solution \mathbf{f}_λ (33) is expressed as

$$\mathbf{f}_\lambda = \mathbf{Y} (\mathbf{\Sigma}^2 + \lambda^2 \mathbf{M}^2)^{-1} \mathbf{\Sigma} \mathbf{U}^H \mathbf{b}_\Pi. \quad (39)$$

- For the WLS with Fixed SNR Loss requirement (36), the solution would occur on the boundary $\|\mathbf{L} \mathbf{f}\|^2 = \alpha$ [26]. Substitute (39) into the boundary, we have

$$\|\mathbf{V} \mathbf{M} \mathbf{Y}^{-1} \cdot \mathbf{Y} (\mathbf{\Sigma}^2 + \lambda^2 \mathbf{M}^2)^{-1} \mathbf{\Sigma} \mathbf{U}^H \mathbf{b}_\Pi\| = \alpha.$$

After some tedious but straightforward manipulations, the relationship between λ and α is given by

$$\sum_{i=1}^n \frac{(\sigma_i \mu_i)^2}{(\sigma_i^2 + \lambda^2 \mu_i^2)^2} \mathbf{b}_\Pi^*(i)^2 = \alpha^2 \quad (40)$$

where $\mathbf{b}_\Pi^* = \mathbf{U}^H \mathbf{b}_\Pi$. From (40), it can be seen that once a threshold value α is specified, there exists a corresponding λ in the regularization framework.

- For the WLS with Fixed Fidelity Loss requirement (37), the solution is also found to be on the boundary $\|\mathbf{G}_\Pi \mathbf{f} - \mathbf{b}_\Pi\|^2 = \beta$ [26]. In the following, we take $\lambda = 1/\tilde{\lambda}$ to simplify the expression of the solution \mathbf{f}_λ . With GSVD of $(\mathbf{G}_\Pi, \mathbf{L})$, we have

$$\mathbf{f}_\lambda = \mathbf{Y}^{-1} (\tilde{\lambda}^2 \mathbf{\Sigma}^2 + \mathbf{M}^2)^{-1} \tilde{\lambda}^2 \mathbf{\Sigma} \mathbf{U}^H \mathbf{b}_\Pi. \quad (41)$$

After substituting (41) back to the boundary condition, it is readily obtained that

$$\|\mathbf{U}^H \mathbf{\Sigma} \mathbf{Y} \cdot \mathbf{Y}^{-1} (\tilde{\lambda}^2 \mathbf{\Sigma}^2 + \mathbf{M}^2)^{-1} \tilde{\lambda}^2 \mathbf{\Sigma} \mathbf{U}^H \mathbf{b}_\Pi - \mathbf{b}_\Pi\| = \beta.$$

Similarly it can be simplified as

$$\sum_{i=1}^n \left(\frac{\mu_i^2 \mathbf{b}_\Pi^*(i)}{\tilde{\lambda}^2 \sigma_i^2 + \mu_i^2} \right)^2 = \beta^2. \quad (42)$$

It is clear that WLS with a quadratic constraint method is a special case of the proposed framework when the constraint values α, β are optimally chosen according to (40) and (42). Although WLS with a quadratic constraint methods are equivalent to the regularization method under some circumstances, generally there is no guideline for choosing optimal α and β . Furthermore, because the elements of \mathbf{G}_Π depend on different timing offset and channel realizations, there is no single λ that is in general optimized for all the scenarios. Thus, fixing the values of α and β in LS with a quadratic constraint as in [26] can fail substantially (more details can be found in the next section).

VI. NUMERICAL RESULTS AND DISCUSSIONS

In this section, the performance of the proposed algorithm for estimation and resynchronization is demonstrated by Monte Carlo simulations, where each point is obtained by averaging over 10^4 runs. In all simulations, the QPSK modulation is used. During the data transmission period, Alamouti Scheme [34] is employed for the $K = 2$ case, and the rate-1/2 orthogonal STBC for four antennas [19] is used for the $K = 4$ case to cooperatively transmit data. In both cases, it can be easily shown that the

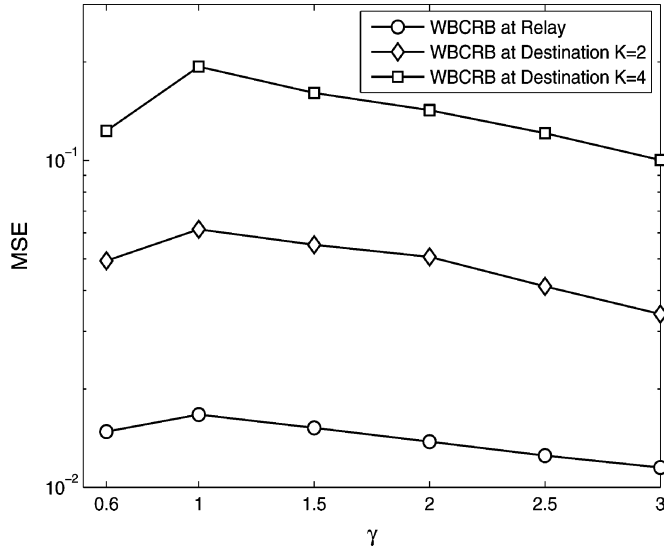


Fig. 5. WBCRB against different values of weighting index γ at SNR = -10 dB.

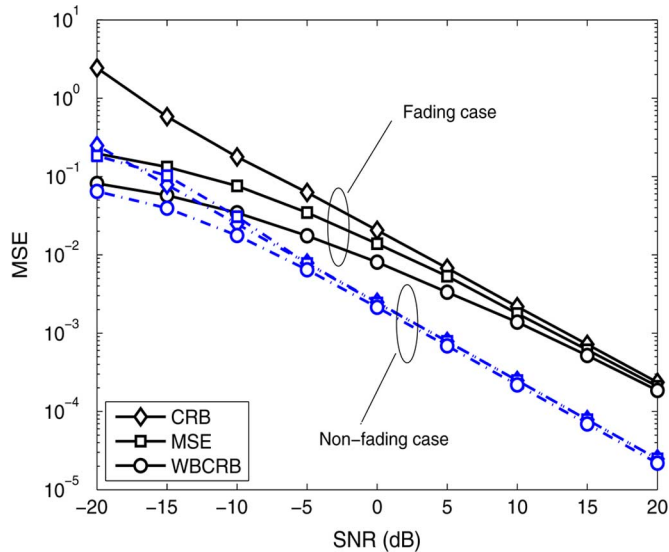


Fig. 6. MSE, CRB, and WBCRB of single timing offset estimation in fading and nonfading environment.

correlation matrix is diagonal as $\mathbf{R}_x = E_x \mathbf{I}$ and the weighting matrix becomes

$$\mathbf{\Pi} \triangleq E_x \begin{bmatrix} |\hat{h}_1|^2 \mathbf{I} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & |\hat{h}_K|^2 \mathbf{I} \end{bmatrix}. \quad (43)$$

The pulse shaping filter $g(t)$ is assumed as root-raised cosine waveform with roll-off factor 0.22 and normalized energy $\int_{-\infty}^{\infty} g^2(t) dt = 1$. In the training period, the training sequence from the source and the relays are all generated as $[\exp(-j\phi_{-L_g}), \dots, \exp(-j\phi_{L_o+L_g})]$, where ϕ_i is uniformly distributed between $[-\pi, \pi]$. The regularization matrix \mathbf{L} is taken to be \mathbf{I} such that $\|\mathbf{L}\mathbf{f}\|^2$ represents the energy of the filter. The channel coefficients are modeled as independent identically distributed (i.i.d.) complex Gaussian random variables with zero mean and unit variance. The SNR is defined

as the average transmit SNR $\text{SNR}_k = E_s/\sigma_{w_k}^2$ for the broadcasting phase and $\text{SNR}_d = E_x/\sigma_v^2$ for the multiple access phase. The length of the training sequence is fixed to $L_o = 65$ and $L_g = 4$.

A. Weighting Index for Weighted Bayesian CRB

Fig. 5 shows the WBCRBs for timing as a function of γ , at both the relay and the destination (with $K = 2$ and $K = 4$). Although numerous values of γ are evaluated for different WBCRB during the simulations, only representative values of $\gamma = [0.6, 1, 1.5, 2, 2.5, 3]$ at SNR = -10 dB are shown in the graph for a clear presentation. We only show the simulation results at low SNR since at high SNR, WBCRBs with different γ all asymptotically converge to the CRB. From Fig. 5, it is noticed that $\gamma = 1$ gives the tightest bound in all cases, and thus we will use $\gamma = 1$ for the rest of the simulations.

B. Estimation Performance

In Fig. 6, we present an illustration of the relationship between CRB, WBCRB, and the MSE of the timing estimate in both fading and nonfading channels at the relay node. The non-fading channel is generated as unit gain with random phase shift $c_k = e^{j\phi_k}$. As can be seen from the curves, CRB is generally a valid bound in nonfading channels (inapplicable in low SNR < -20 dB). However in fading channels, the validity of CRB does not hold until the average SNR becomes considerably large (at least >20 dB) because the equivalent (receive) SNR of a fading channel is comparatively smaller, which levers up the equivalent SNR threshold for the estimate to touch the CRB. On the other hand, by incorporating the prior knowledge on timing, the WBCRB is valid for both cases even under deep fading because the prior knowledge limits the error within a certain range. Notice that at low SNR, the WBCRB is reduced to the variance of the uniform distribution of timing offsets (i.e., $\text{Var}\{\tau_k\} = 1/12$), while at high SNR, the WBCRB asymptotically converges to the CRB.

In Fig. 7, the CRB, the WBCRB and the corresponding timing estimation MSE at the destination node are plotted as a function of SNR. The timing MSE is defined as the average MSE over all the simulation runs $\text{MSE}(\epsilon) = \sum_{n=1}^N \sum_{k=1}^K (\hat{\epsilon}_k^{(n)} - \epsilon_k)^2 / N$, where $\hat{\epsilon}_k^{(n)}$ is the estimate of the k th offset in the n th simulation run, with $N = 10^4$ being the total number of simulation trials. It can be seen that the CRB and WBCRB coincide in high SNR region while the CRB becomes inapplicable even when the SNR is relatively high (30 dB). This is caused by the fading as well as the fact that the CRB simply considers the parameter as a real number with infinite range, hence it goes upward infinitely as SNR decreases. At low SNR, the WBCRB is reduced to the variance of the uniform distribution of timing offsets, i.e., $\sum_{k=1}^K \text{Var}\{\epsilon_k\} = K/12$.

The MSE of the channels $\text{MSE}(\mathbf{h}) = \sum_{n=1}^N \sum_{k=1}^K (\hat{h}_k^{(n)} - h_k)^2 / N$ and the corresponding CRB are plotted in Fig. 8, where $\hat{h}_k^{(n)}$ is the estimate of the k th channel coefficient in the n th simulation run. In Fig. 8, it is clear that on the contrary to the timing estimates, the channel MSEs are well bounded by the CRB since the values of the channel taps do not assume any *a priori* given range as in the case of the timing offsets. Finally, it can be seen

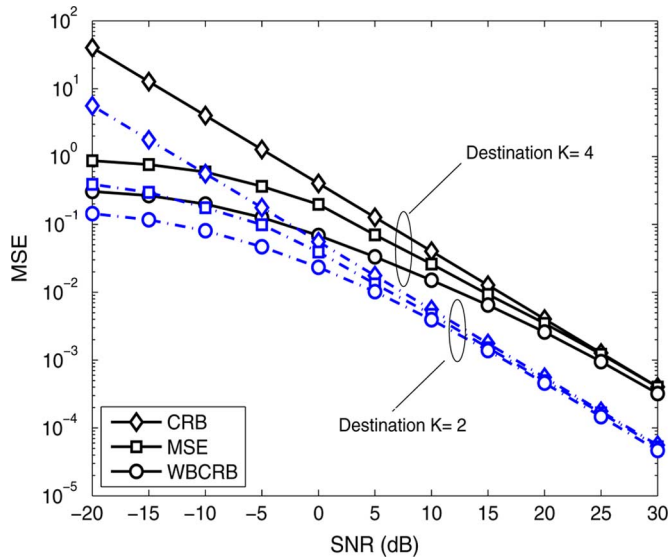


Fig. 7. MSE, CRB, and WBCRB for timing estimation at the destination ($K = 2$ and 4) with fading channels.

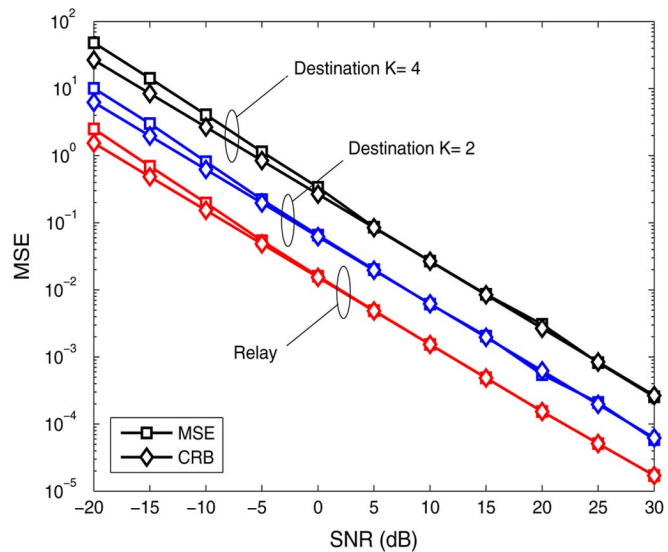


Fig. 8. MSE and CRB for channel estimation at the relay and destination with fading channels.

from Fig. 6 to 8 that the timing and channel estimators are efficient under high SNR since the MSE curves touch the CRB for all the cases as SNR increases.

C. Synchronization Performance

The performances of the resynchronization algorithms proposed in this paper are thoroughly illustrated in this section.

In Fig. 9, the performance of the proposed regularized filter is illustrated against several LS methods with $K = 2$ and travel delay differences $|\Delta_k| < 0.1T$. It can be seen that the performance of the regularized filter overlaps with the ideal case (i.e., $\Delta_k = 0$) while the ordinary LS solution provides disappointing results due to the ill-conditioned nature of the problem. On the other hand, for the LS with a quadratic constraint [26], it is obvious that the nonoptimal quadratic constraint parameters α lead to significant performance degradation, especially in the case with $\alpha = 4$. Similar results are obtained for the Fixed Fidelity

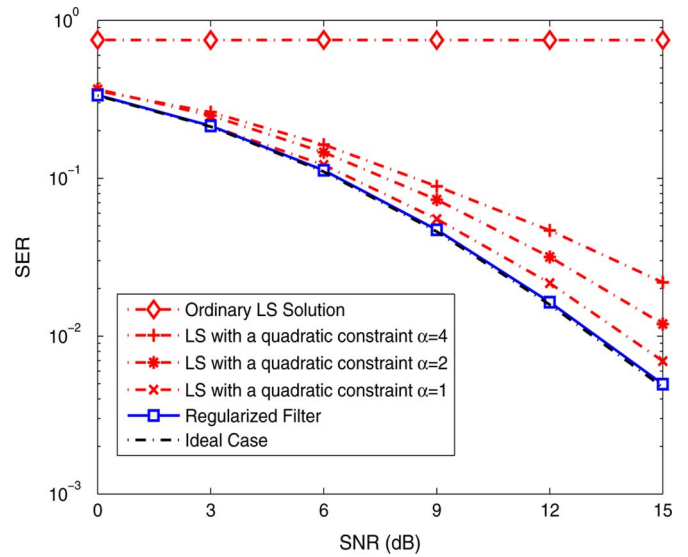


Fig. 9. Symbol error rate (SER) performance for resynchronization filter designed by Weighted Regularized LS against the performance by ordinary LS solution, and LS with a quadratic constraint with $\alpha = 4, 2, 1$ under QPSK modulation, $|\Delta_k| < 0.1T$ for $K = 2$.

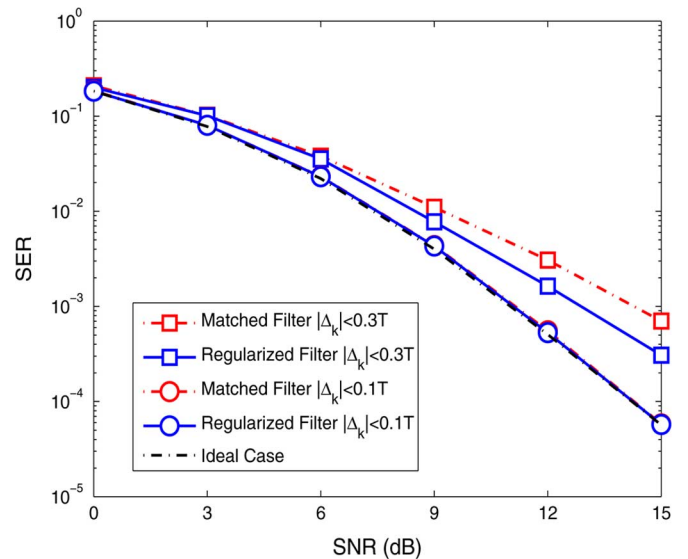


Fig. 10. SER performance of the Regularized Filter and the Matched Filter under QPSK modulation with travel delay differences $|\Delta_k| < 0.1$ and $0.3T$, respectively, for $K = 4$.

Loss method, and, thus, not presented here. The performances of weighted LS with a quadratic constraint are not displayed because although they are slightly better than the LS method in [26], no significant gap is found.

In Fig. 10, the regularized filter is compared against the matched filter under different travel delay differences of $|\Delta_k| < 0.1$ and $0.3T$ (with $K = 4$). With the data correlation in (43), the matched filter is the weighted mean delay filter with nominal delay $\epsilon_c = \sum_{k=1}^K |\hat{h}_k|^2 \hat{\epsilon}_k / \sum_{k=1}^K |\hat{h}_k|^2$. It is observed that when the delay difference is relatively small ($\Delta_k < 0.1T$), the performance of the regularized filter and the matched filter both overlap with the ideal case. For medium to large timing jitters ($|\Delta_k| < 0.3T$), the performance of the matched filter degrades considerably due to the increasing residual ISI which

results from the uncompensated timing mismatch as well as the filter shape constraint. On the other hand, the performance of the regularized filter under large timing jitters remains at a satisfactory level. Although the matched filter is less accurate, it is a good engineering tradeoff in terms of computational and implementation complexity.

Next, we provide a simple calculation on the relationship between the actual distance and the timing delay difference Δ_k , which is normalized to the symbol duration T . The travel distance differences can be calculated accordingly as $\Delta l \in (-|\Delta_k|cT, |\Delta_k|cT)$, where c is the speed of light. Generally, the delay spread of a wireless channel is on the order of 10^{-6} s [35], so for a single carrier system over flat fading channels, the symbol duration T is usually on the order of 10^{-5} s. Hence the distance difference can amount to $(-300 \text{ m}, 300 \text{ m})$ for $|\Delta_k| < 0.1T$ and $(-900 \text{ m}, 900 \text{ m})$ for $|\Delta_k| < 0.3T$. Therefore, the proposed filter can generally perform well even when deploying a cluster of relays that are quite far away from each other.

VII. CONCLUSION

In this paper, the problem of joint timing offset and channel estimation, and furthermore the resynchronization of multiple timing offsets in a decode and forward cooperative relay system has been considered. The ML estimators, CRB and WBCRB are derived for the timing and channel parameters. In order to design the resynchronization filter, two methods have been proposed. The first one is a matched filter that minimizes the residual ISI power. Its implementation is simple and provides satisfactory performance. The second one is the weighted regularized least squares method. The proposed regularized filter is novel and effective in that: 1) The design parameters in the proposed framework are optimally chosen for each realization of channels and offsets; 2) The weighting incorporated into the regularization effectively extracts the information from those less attenuated components in the received signal. The regularized resynchronization filter outperforms the matched filter at the expense of higher computational complexity. Simulation results have well supported our presented analysis and also verified the efficiency of the estimation and synchronization algorithms.

APPENDIX

PROOF OF $A_\Sigma > 0$ IN (29)

For notation simplicity, define $\mathbf{d}_{gg} \triangleq [R'_{gg}(-L_o - M_o), \dots, R'_{gg}(0), \dots, R'_{gg}(L_o + M_o)]^T$. Then using matrix notation, we can obtain $f(k_1, k_2) = \mathbf{d}_{gg}^H \mathbf{R}_{k_1, k_2} \mathbf{d}_{gg}$. Furthermore, denote matrix \mathcal{F} such that $[\mathcal{F}]_{k_1, k_2} = f(k_1, k_2)$, then \mathcal{F} can be written compactly as

$$\mathcal{F} = \begin{bmatrix} \mathbf{d}_{gg} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{d}_{gg} \end{bmatrix}^H \times \underbrace{\begin{bmatrix} \mathbf{R}_{1,1} & \cdots & \mathbf{R}_{1,K} \\ \vdots & \ddots & \vdots \\ \mathbf{R}_{K,1} & \cdots & \mathbf{R}_{K,K} \end{bmatrix}}_{=\mathbf{R}_x} \underbrace{\begin{bmatrix} \mathbf{d}_{gg} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{d}_{gg} \end{bmatrix}}_{\triangleq \mathbf{D}_{gg}}. \quad (44)$$

Thus, A_Σ can be reexpressed as

$$A_\Sigma = \hat{\mathbf{h}}^H \mathcal{F} \hat{\mathbf{h}} \quad (45)$$

where $\hat{\mathbf{h}} = [\hat{h}_1, \dots, \hat{h}_K]^T$. Therefore, $A_\Sigma > 0$ is equivalent to matrix \mathcal{F} being positive definite.

In order for $\mathcal{F} = \mathbf{D}_{gg}^H \mathbf{R}_x \mathbf{D}_{gg}$ to be positive definite, it is required that $\mathbf{y}^H \mathbf{D}_{gg}^H \mathbf{R}_x \mathbf{D}_{gg} \mathbf{y} > 0$ for any vector \mathbf{y} . Since $\mathbf{R}_x = \mathbb{E}\{\mathcal{X}\mathcal{X}^H\}$, we have

$$\begin{aligned} \mathbf{y}^H \mathbf{D}_{gg}^H \mathbf{R}_x \mathbf{D}_{gg} \mathbf{y} &= \mathbb{E}\{\mathbf{y}^H \mathbf{D}_{gg}^H \mathcal{X} \mathcal{X}^H \mathbf{D}_{gg} \mathbf{y}\} \\ &= \mathbb{E}\{|\mathcal{X}^H \mathbf{D}_{gg} \mathbf{y}|^2\} \\ &= \underbrace{\text{Var}\{|\mathcal{X}^H \mathbf{D}_{gg} \mathbf{y}|\}}_{>0} + \underbrace{\mathbb{E}^2\{|\mathcal{X}^H \mathbf{D}_{gg} \mathbf{y}|\}}_{\geq 0} > 0 \end{aligned}$$

where the term $|\mathcal{X}^H \mathbf{D}_{gg} \mathbf{y}|$ is the magnitude of the weighted conjugated sum of the elements in vector \mathcal{X} . The first inequality $\text{Var}\{|\mathcal{X}^H \mathbf{D}_{gg} \mathbf{y}|\} > 0$ comes from the fact that \mathcal{X} is a random vector (data), and the variance of a random variable is always positive (since it is equal to zero if and only if \mathcal{X} is constant or exhibits extreme correlation characteristics).

Now that $\mathcal{F} = \mathbf{D}_{gg}^H \mathbf{R}_x \mathbf{D}_{gg}$ is a positive definite Hermitian matrix, it is guaranteed that

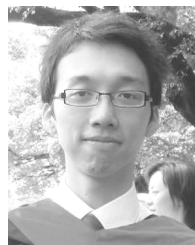
$$A_\Sigma = \hat{\mathbf{h}}^H \mathcal{F} \hat{\mathbf{h}} > 0. \quad (46)$$

REFERENCES

- [1] D. Gesbert, M. Shafi, D.-S. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 281–302, 2003.
- [2] A. Naguib, N. Seshadri, and R. Calderbank, "Increasing data rate over wireless channels," *IEEE Signal Process. Mag.*, vol. 17, no. 76–92, May 2000.
- [3] G. J. Foschini, G. D. Golden, P. W. Wolniansky, and R. A. Valenzuela, "Simplified processing for wireless communication at high spectral efficiency," *IEEE J. Sel. Areas Commun.-Wireless Commun. Ser.*, vol. 17, pp. 1841–1852, 1999.
- [4] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity, part I, II," *IEEE Trans. Commun.*, vol. 51, pp. 1927–1948, Nov. 2003.
- [5] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.
- [6] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, pp. 1089–1098, Aug. 2004.
- [7] P. A. Anghel and M. Kaveh, "Exact symbol error probability of a cooperative network in a rayleigh-fading environment," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1416–1421, Sept. 2004.
- [8] A. Ribeiro, X. Cai, and G. B. Giannakis, "Symbol error probabilities for general cooperative links," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1264–1273, May 2005.
- [9] R. U. Nabar and H. Bolcskei, "Space-time signal design for fading relay channels," in *Proc. GlobeCom*, Dec. 2003, pp. 1952–1956.
- [10] O.-S. Shin, A. M. Chan, H. T. Kung, and V. Tarokh, "Design of an OFDM cooperative space-time diversity system," *IEEE Trans. Commun.*, vol. 56, no. 4, pp. 2203–2215, Jul. 2007.
- [11] J. Mietzner and P. A. Hoeher, "Distributed space-time codes for cooperative wireless networks in the presence of different propagation delays and path losses," in *Proc. 3rd IEEE Sens. Array Multichannel Signal Process. Workshop*, Barcelona, Spain, Jul. 2004, pp. 264–268.
- [12] R. C. Palat, A. Annamalai, and J. H. Reed, "Upper bound on bit error rate for time synchronization errors in bandlimited distributed MIMO networks," in *Proc. IEEE WCNC*, Apr. 2006, vol. 4, pp. 2058–2063.

- [13] Y. Mei, Y. Hua, A. Swami, and B. Daneshrad, "Combating synchronization errors in cooperative relays," in *Proc. IEEE ICASSP*, Mar. 2005, vol. 3, pp. 369–372.
- [14] S. Jagannathan, H. Aghajan, and A. Goldsmith, "The effect of time synchronization errors on the performance of cooperative MISO systems," in *Proc. IEEE Globecom*, Nov. 2004, pp. 102–107.
- [15] R. A. Iltis and R. Cagley, "Channel estimation and carrier offset control for cooperative MIMO sensor networks," in *IEEE Proc. 39th Asilomar Conf. Signals, Systems, Comput.*, Pacific Grove, CA, Nov. 2005, vol. 1, pp. 210–214.
- [16] B. Azimi-Sadjadi and A. Mercado, "Diversity gain for cooperating nodes in multi-hop wireless networks," in *Proc. IEEE VTC-Fall*, Sep. 2004, vol. 2, pp. 1483–1487.
- [17] R. Djapic, A.-J. Van der Veen, and L. Tong, "Synchronization and packet separation in wireless ad-hoc networks by known modulus algorithm," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 1, pp. 51–64, Jan. 2005.
- [18] P. Stoica and E. Lindskog, "Space-time block coding for channels with intersymbol interference," in *IEEE Proc. 35th Asilomar Conf. Singlas, Syst., Comput.*, Pacific Grove, CA, Nov. 2001, vol. 1, pp. 252–256.
- [19] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
- [20] X. Li, "Space-time coded multi-transmission among distributed transmitters without perfect synchronization," *IEEE Signal Process. Lett.*, vol. 11, no. 12, pp. 948–951, Jan. 2005.
- [21] T. Wang, Y. Yao, and G. B. Giannakis, "Non-coherent distributed space-time processing for multiuser cooperative transmission," in *Proc. IEEE GlobeCom*, Nov. 2005, pp. 3738–3742.
- [22] P. A. Parker, D. W. Bliss, P. Mitran, and V. Tarokh, "Adaptive frequency synchronization for collaborative communication systems," in *Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Jun. 2007.
- [23] K. Kim, H. Kim, and H. Park, "OFDM channel estimation for the amplify-and-forward cooperative channel," in *Proc. IEEE Veh. Technol. Conf.*, Apr. 2007.
- [24] Y.-C. Wu and E. Serpedin, "Design and analysis of feedforward symbol timing estimators based on the conditional maximum likelihood principle," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1908–1918, May 2005.
- [25] A. Hafeez, K. J. Molnar, H. Arslan, G. E. Bottomley, and R. Ramesh, "Adaptive joint detection of cochannel signals for TDMA handsets," *IEEE Trans. Commun.*, vol. 52, pp. 1722–1732, Oct. 2004.
- [26] A. Kannan, T. P. Krauss, and M. D. Zoltowski, "Separation of co-channel signals under imperfect timing and carrier synchronization," *IEEE Trans. Veh. Technol.*, vol. 50, pp. 79–96, Jan. 2001.
- [27] J. Riba, J. Sala, and G. Vazquez, "Conditional maximum likelihood timing recovery: Estimators and bounds," *IEEE Trans. Signal Process.*, vol. 49, pp. 835–850, Apr. 2001.
- [28] M.-O. Pun, M. Morelli, and C.-C. Jay Kuo, "Maximum-likelihood synchronization and channel estimation for OFDMA uplink transmissions," *IEEE Trans. Commun.*, vol. 54, no. 4, pp. 726–736, Apr. 2006.
- [29] 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.
- [30] H. L. Van Trees and K. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. Piscataway, NJ: Wiley-IEEE, 2007.
- [31] J. Ziv and M. Zakai, "Some lower bounds on signal parameter estimation," *IEEE Trans. Inf. Theory*, vol. 15, no. 3, pp. 386–391, May 1969.
- [32] P. C. Hansen, *Rank Deficient and Discrete Ill-Posed Problems-Numerical Aspects of Linear Inversion*. Philadelphia, PA: SIAM, 1998.
- [33] A. H. Sayed, *Fundamentals of Adaptive Filtering*. Piscataway, NJ: Wiley-IEEE, 2003.

- [34] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.
- [35] H. Bölcskei, "Delay spread of several measured channels, and physical-layer parameters of common wireless systems," Communication Technology Laboratory, ETH [Online]. Available: <http://www.nari.ee.ethz.ch/commth/teaching/wirelessIT/handouts/phy-params-handout.pdf> May 2006, [Online]. Available



Xiao Li received the B.Eng. degree (*summa cum laude*) from Sun Yat-sen (Zhongshan) University in 2007. He is currently pursuing the M.Phil. degree at the Department of Electrical and Electronic Engineering, the University of Hong Kong.

His current research interests include statistical signal processing, optimization, regularization theory and sparse signal recovery as well as Bayesian analysis and Monte Carlo methods, with applications in MIMO, OFDM, and cooperative communication systems.



Yik-Chung Wu received the B.Eng. (EEE) degree in 1998 and the M.Phil. degree in 2001 from The University of Hong Kong (HKU) and the Ph.D. degree in 2005 from Texas A&M University.

From August 2005 to August 2006, he was with the Thomson Corporate Research, Princeton, NJ, as a Member of Technical Staff. Since September 2006, he has been with the University of Hong Kong as an Assistant Professor. His research interests are in the general area of signal processing and communication systems, and in particular, receiver algorithm design, synchronization techniques, channel estimation and equalization.

Dr. Wu is currently serving as an Associate Editor for the IEEE COMMUNICATIONS LETTERS.



Erchin Serpedin received (with highest distinction) the Diploma of Electrical Engineering from the Polytechnic Institute of Bucharest, Bucharest, Romania, in 1991. He received the specialization degree in signal processing and transmission of information from Ecole Supérieure D'Electricité, Paris, France, in 1992, the M.Sc. degree from the Georgia Institute of Technology, Atlanta, in 1992, and the Ph.D. degree in electrical engineering from the University of Virginia, Charlottesville, in January 1999.

In July 1999, he joined the Wireless Communications Laboratory, Texas A&M University, College Station, as an Assistant Professor, and currently holds the position of Associate Professor. His research interests lie in the area of signal processing and its applications in wireless communications, bioinformatics, and genomics. He is the author of 60 journal papers, 80 conference papers, and of the upcoming research monograph *Synchronization in Wireless Sensor Networks* (Cambridge, U.K.: Cambridge University Press, 2009).