EURASIP Journal on Applied Signal Processing 2005:3, 273–286


# Multiuser Channel Estimation for Ultra-Wideband Systems Using up to the Second-Order Statistics

**Zhengyuan Xu**
*Department of Electrical Engineering, University of California, Riverside, CA 92521, USA*
*Email: dxu@ee.ucr.edu*

**Jin Tang**
*Department of Electrical Engineering, University of California, Riverside, CA 92521, USA*
*Email: jtang@ee.ucr.edu*

**Ping Liu**
*Department of Electrical Engineering, Arkansas Tech University, Russellville, AR 72801, USA*
*Email: ping.liu@mail.atu.edu*

*Received 27 September 2003; Revised 11 March 2004*

In a pulse-position modulation-based ultra-wideband (UWB) communication system, multiple access is enabled by assigning unique time-hopping sequences to different users. Each user's data information is carried by positions of short pulses which are directly transmitted through an unknown and possibly dense multipath channel. Single-user channel estimation methods have been proposed by maximum likelihood optimization that treats multiple access interference as Gaussian noise. In this paper, multiuser channel estimation methods are proposed based on a pulse-rate discrete-time system model and up to the second-order statistics of the channel outputs. The model can be regarded in a trilinear structure and also resembles a code-division multiple-access (CDMA) system with newly defined hopping-code dependent matrices and inputs for each user. Considering that either the mean or covariance of received signals contains sufficient information for all unknown channels, least squares and covariance matching ideas are successfully applied to estimate all channels blindly. Accordingly, closed-form solutions are derived. Those channel estimates can be used to design typical linear receivers. Performance of each proposed estimator is analyzed and also verified by computer simulations. Corresponding receivers' performance is also studied numerically.

**Keywords and phrases:** ultra-wideband, channel estimation, covariance matching, least squares.

## 1. INTRODUCTION

Ultra-wideband (UWB) technology is originated from works in the time-domain electromagnetics early in the 1960s [1, 2]. It is based on a widely recognized fact that electromagnetic signals for radio transmission and radar do not need to have an approximately sinusoidal time variation, as discussed in detail in [3] which shows that waves with arbitrary time variation can be radiated. With a novel antenna design technique [4], the generated UWB signals are able to communicate by baseband short pulses. Thereafter, the technique was immediately applied to radar, communications, automobile collision avoidance, positioning systems, and liquid-level sensing [5]. The term "ultra-wideband" was not applied until late 1980s by the US Department of Defense when a covert property with low probability of interception and detection (LPI/LPD) was realized.

With a recent release of the spectral mask from the Federal Communications Commission (FCC) [6], there emerges an increasing interest in UWB techniques for both commercial and military applications [7, 8]. In general, a UWB signal is defined as having 10 dB bandwidth greater than 20% of its center frequency [6]. UWB systems transmit trains of time-hopping (TH) short-duration pulses with a low duty cycle and use pulse-position modulation (PPM) [7, 9], and thus can be termed as TH-UWB. Those subnanosecond pulses are often referred to as monocycles. Therefore, a multipath down to path delay differentials in nanoseconds is resolvable at the receiver, significantly mitigating multipath distortion and providing path diversity [10, 11]. Meanwhile, very low power transmission insures little interference incurred to existing narrowband systems operating in the same frequency band. With a properly designed pulse characterized by spectral property [12] and selected hopping codes, multiple access (MA) is maximally enabled for

different users to simultaneously share large bandwidth. The nature of impulse radio (IR) also makes the radio wave easily penetrate materials and obstructs, desirable for non-line-of-sight communication environments. It thus becomes not only an ideal candidate for communications in dense multipath environments such as short-range or indoor wireless communications [13, 14], but also viable in government and military wireless networks for support of both tactical and strategic communications [8].

Successful deployment of a UWB system in a wireless environment requires reliable symbol detection techniques. A conventional UWB receiver is a RAKE receiver. It has a very simple structure and consists of waveform correlators [7]. To fully capture the signal energy spread over multiple paths, the receiver needs to know channel parameters when correlation is performed. Channel parameters are also required by some existing methods in design of multiuser receivers [15, 16]. However, in a dense multipath wireless environment, channel information can not be known *a priori*. Although single-user maximum likelihood (ML) channel estimation methods have been proposed [13, 17], multiple access interference (MAI) is approximated as a Gaussian process which may not be accurate and has to be considered explicitly for performance improvement. Although a low complexity channel estimation method has been proposed in [18], it is designed for peer-to-peer communications and requires training sequences.

In this paper, we study blind multiuser channel estimation from statistics of received signals in an MA-UWB system. For low complexity and easy implementation, only first-order and second-order statistics (SOS) are employed. First, we adopt a pulse-rate discrete-time channel model developed in [15] which makes blind channel estimation possible. It is then observed that a UWB system resembles a direct-sequence (DS) code-division multiple-access (CDMA) system. After clearly defining a matrix for each user from its unique TH sequence, each matrix can be treated as a code matrix, similar to the code matrix constructed from spreading codes in a CDMA system [19]. But it is sparse and consists of only zeros and ones, indicating whether there exists a contribution to the received signal during a particular time interval from a multipath channel or not. Locations of zeros and ones are different for different users, capable of differentiating users. After such linear modeling, PPM is transformed to superimposed amplitude modulation that is easy to handle. The received signal is examined to exhibit nonzero mean that is linearly parameterized by multiuser channels. Therefore, our first channel estimation approach is based on a least-squares (LS) criterion that minimizes the error between the estimated mean from data and its model-based one. All channels are then estimated without any ambiguity. Secondly, assisted by unique code matrices, a covariance matching (CM) idea can be applied to estimate each channel. Data covariance is parameterized by a rank-one channel-dependent matrix of each user. If the CM error is minimized, then each channel can be estimated up to a phase ambiguity [19]. Here covariance instead of correlation is used since nonzero mean incurs channel cross products in the auto-correlation of directly received data and complicates estimation. Although chip-rate sampling induced multiple-input multiple-output (MIMO) model can be used for symbol detection for given channel parameters [15], it creates multiple subchannels corresponding to the same code matrix and same propagation channel for each user. Under that modeling, LS method can only yield an estimate of linearly superimposed subchannels, while CM method can only provide an estimate with a unitary matrix ambiguity. Therefore, only pulse-rate sampling is adopted in this paper.

There is no doubt that TH-UWB is not the only signaling format to support UWB MA communications. IEEE is at its early stage of discussions on UWB IEEE 802.15.3a standard (see, e.g., [20]). DS-CDMA appears as an competitive alternative due to its favorable property of narrowband interference rejection [21, 22]. Comparisons between TH-UWB and DS-UWB system performance are made in [23, 24]. However, we only focus on TH-UWB PPM systems in this paper although the CM technique can be generalized to UWB systems with CDMA modulation format like conventional CDMA systems [19].

### *Notations*

Following common practice, we denote the Kronecker product by $\otimes$, Hadamard (elementwise) product by $\odot$, complex conjugate ( $^*$ ) transpose ( $^{\mathrm{T}}$ ) by ( $^{\mathrm{H}}$ ), inverse by ( $^{-1}$ ). $E\{\cdot\}$ represents expectation of a random variable, $\mathbf{I}_a$ an identity matrix of degree $a$ whose $i$th column is denoted by $\mathbf{e}_{a,i}$. $\mathbf{1}_a$ is a vector of length $a$ with all elements equal to one. $\mathrm{vec}(\cdot)$ is an operator to stack all column vectors of a matrix successively in a big column vector. An estimate of a quantity (scalar, vector, or matrix) is denoted by putting a hat "$\hat{}$" over it, and correspondingly, the estimation error by preceding the quantity with a $\delta$, such as $\hat{\mathbf{x}}$ and $\delta\mathbf{x}$ for vector $\mathbf{x}$, respectively. However, $\delta(\cdot)$ represents a discrete-time unit-impulse function. $\lfloor\cdot\rfloor$ stands for integer floor, while $\lceil\cdot\rceil$ for integer ceiling. We define the distributive Kronecker product $\mathbf{X}\diamond\mathbf{Y} = [\mathbf{X}\otimes\mathbf{Y}(:,1), \mathbf{X}\otimes\mathbf{Y}(:,2), \ldots]$ based on columns of $\mathbf{Y}$.

## 2. SYSTEM MODEL

Consider an MA-TH-UWB system with $K$ users. The transmitted baseband UWB signal from user $k$ can be described by [15]

$$\alpha_k(t) = \sqrt{\mathcal{P}_k} \sum_{i=-\infty}^{\infty} w\left(t - iT_f - c_k(i)T_c - \tau_{I_k(\lfloor i/N_f \rfloor)}\right), \quad (1)$$

where $\mathcal{P}_k$ is the $k$th user's transmission power, $w(t)$ is the baseband monopulse, $T_f$ is the frame duration, $N_f$ is the number of frames over which an $M$-ary PPM symbol repeats, $c_k(i) \in [0, N_c - 1]$ is a periodic hopping sequence with the period equal to one symbol period. Each chip has duration $T_c$. $I_k(\lfloor i/N_f \rfloor) \in [0, M-1]$ is the $k$th user's information bearing symbol during the $i$th frame, $\tau_{I_k(\lfloor i/N_f \rfloor)} = I_k(\lfloor i/N_f \rfloor)\sigma$ is the corresponding modulation delay in a multiple of $\sigma$ seconds. Assume $T_f = N_c T_c$ and $T_c = M\sigma$.

If we define $w_m(t) \triangleq w(t - m\sigma)$, where $m = 0, \ldots, M-1$ and $s_{k,m}(\lfloor i/N_f \rfloor) = \delta(I_k(\lfloor i/N_f \rfloor) - m)$, then (1) may be expressed by linear modulation in a chip rate as [15]

$$\alpha_k(t) = \sqrt{\mathcal{P}_k} \sum_{i=-\infty}^{\infty} \sum_{m=0}^{M-1} u_{k,m}(i) w_m(t - iT_c), \tag{2}$$

where the chip index has replaced the frame index in (1),

$$\begin{aligned} u_{k,m}(i) &= s_{k,m}\left(\left\lfloor \frac{i}{N_c N_f} \right\rfloor\right) \widetilde{c}_k(i), \\ \widetilde{c}_k(i) &= \delta\left(\left\lfloor \frac{i}{N_c} \right\rfloor N_c + c_k\left(\left\lfloor \frac{i}{N_c} \right\rfloor\right) - i\right). \end{aligned} \tag{3}$$

It is clear according to (2) that the input $u_{k,m}(i)$ is modulated by waveform $w_m(t)$ at a chip rate. The transmitted signal $\alpha_k(t)$ propagates through a linear channel with impulse response $\bar{g}_k(t)$. At the receiver, the channel output is first passed through a filter matched to the monopulse $w(t)$ [16]. We can define a front-end effective channel including effects from modulated pulse at the transmitter, and propagation channel and matched filter at the receiver by $g_{k,m}(t) = w_m(t) \star \bar{g}_k(t) \star w(-t)$, where $\star$ denotes convolution. Considering additive white Gaussian noise (AWGN) $v(t)$ and propagation delay $d_k$ for user $k$, the output of the matched filter becomes

$$y(t) = \sum_{k,i_1,m} \sqrt{\mathcal{P}_k} u_{k,m}(i_1) g_{k,m}(t - i_1 T_c - d_k) + v(t). \tag{4}$$

Then $y(t)$ is sampled every $\sigma$ seconds to yield a discrete-time output $y(n) = y(t)|_{t=n\sigma}$. Using the discrete-time version of the effective channel and invoking $T_c = M\sigma$, we obtain a pulse-rate model

$$y(n) = \sum_{k,m} \sum_{i_2=1}^{q} \sqrt{\mathcal{P}_k} u_{k,m}\left(\frac{n - i_2}{M}\right) g_{k,m}(i_2) + v(n), \tag{5}$$

where we have assumed the maximum channel length of all users is $q\sigma$. Coefficients of each discrete-time channel in (5) are obtained from sampling the corresponding continuous time channel in (4) and absorbing effect of associated propagation delay. Therefore, each discrete-time channel may contain certain number of leading and/or trailing zeros due to asynchronous transmission from different users. In a synchronous situation, sampling is performed at the beginning of each symbol period. Then there are no leading zeros for each channel. The number of possible trailing zeros is equal to the number of samples within time difference between $q\sigma$ and the individual channel delay spread. In an asynchronous situation, sampling is performed at the beginning of each symbol period corresponding to the user with the minimum propagation delay ($d_{\min} = \min_k d_k$). No leading zeros occur for that user's channel. The number of trailing zeros is the same as discussed above. For each of all other users' channels, the number of leading zeros is equal to the number of samples within time $d_k - d_{\min}$, while the number of trailing zeros is equal to the number of samples within time difference between $q\sigma - (d_k - d_{\min})$ and the corresponding channel delay spread. Consider $P$ symbol intervals of data samples with corresponding time instants $nMN_cN_f + p$ for $p = 1, \ldots, MPN_cN_f$ and collect them in a big vector $\mathbf{y}_n$ of length $\nu = MPN_cN_f$. After noticing our definition of $u_{k,m}(i)$, a vector form data model follows:

$$\mathbf{y}_n = \sum_{k,m,l} \mathbf{C}_{k,l} \mathbf{T}_m \mathbf{g}_k s_{k,m}(n + l) + \mathbf{v}_n, \tag{6}$$

where the symbol index $l$ takes all integers $-\lceil q/(MN_cN_f) \rceil, \ldots, P-1$, $\mathbf{g}_k$ is an unknown channel vector for user $k$ which contains channel coefficients at the pulse rate and power factor $\sqrt{\mathcal{P}_k}$, $\mathbf{T}_m = [\mathbf{0}, \mathbf{I}_q, \mathbf{0}]^{\mathrm{T}}$ is a tall selection matrix in order to obtain the $m$th subchannel from $\mathbf{g}_k$ (delayed in $m\sigma$ seconds or equivalently downshifted by $m$ elements), and $\mathbf{C}_{k,l}$ is a matrix constructed from corresponding $\widetilde{c}_k(i)$ and is uniquely determined by the TH sequence. It consists of only zeros and ones and repeats from symbol to symbol because the TH sequence has a period equal to one symbol interval. This model can be compactly expressed in another form:

$$\mathbf{y}_n = \sum_{k,l} \mathbf{H}_{k,l} \mathbf{s}_{k,n,l} + \mathbf{v}_n = \mathbf{H}\mathbf{s}_n + \mathbf{v}_n, \tag{7}$$

after collecting $M$ inputs in a vector

$$\mathbf{s}_{k,n,l} = \left[s_{k,0}(n + l), \ldots, s_{k,M-1}(n + l)\right]^{\mathrm{T}}, \tag{8}$$

defining a corresponding effective channel matrix

$$\mathbf{H}_{k,l} = \left[\mathbf{C}_{k,l}\mathbf{T}_0\mathbf{g}_k, \ldots, \mathbf{C}_{k,l}\mathbf{T}_{M-1}\mathbf{g}_k\right], \tag{9}$$

and successively stacking such matrices (or vectors) in $\mathbf{H}$ (or $\mathbf{s}_n$). The total number of symbols from $K$ users is denoted by $L = K(P + \lceil q/(MN_cN_f) \rceil)$. By employing data model (6), all channels can be estimated based on the statistics (mean or covariance) of $\mathbf{y}_n$.

## 3. BLIND CHANNEL ESTIMATION

It is observed that all channel vectors are embedded in the statistics of the data vector. For low complexity, we consider either its mean or covariance, yielding LS or CM method accordingly.

### 3.1. LS approach

We denote the mean of $\mathbf{y}_n$ as $\bar{\mathbf{y}}$. From our definition, the mean of $\mathbf{s}_{k,n,l}$ is easily found to be $(1/M)\mathbf{1}_M$. Since noise has zero mean even after the matched filter, we have

$$\bar{\mathbf{y}} = \frac{1}{M} \sum_{k,m,l} \mathbf{C}_{k,l} \mathbf{T}_m \mathbf{g}_k = \sum_k \mathcal{C}_k \mathbf{g}_k = \mathcal{C}\mathbf{g}, \tag{10}$$

where all channel vectors are stacked in a big vector $\mathbf{g}$. Assume $\bar{\mathbf{y}}$ is estimated from $N$ data vectors by sample average

$$\hat{\bar{\mathbf{y}}} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{y}_n. \tag{11}$$

Then an LS criterion can be applied to estimate $\mathbf{g}$ as follows:

$$\hat{\mathbf{g}} = \arg\min ||\bar{\mathbf{y}} - \hat{\bar{\mathbf{y}}}||^2. \tag{12}$$

Invoking (10), the solution to (12) has the following form:

$$\hat{\mathbf{g}} = \mathbf{W}\hat{\bar{\mathbf{y}}}, \qquad \mathbf{W} = (\mathcal{C}^{\mathrm{H}}\mathcal{C})^{-1}\mathcal{C}^{\mathrm{H}}. \tag{13}$$

Then the estimate of $\mathbf{g}_k$ can be obtained from the corresponding subvector of $\hat{\mathbf{g}}$ as $\hat{\mathbf{g}}_k = (\mathbf{e}_{K,k}^{\mathrm{T}} \otimes \mathbf{I}_q)\hat{\mathbf{g}}$.

### 3.2. CM approach

Since $\mathbf{y}_n$ has nonzero mean, its autocorrelation is found to have cross terms $\mathbf{g}_{k_1}\mathbf{g}_{k_2}^{\mathrm{H}}$ of users $k_1$ and $k_2$ and not convenient for channel estimation. Thus covariance is considered. Define a new zero-mean data vector from $\mathbf{y}_n$ as

$$\mathbf{z}_n = \mathbf{y}_n - \bar{\mathbf{y}} = \sum_{k,l} \mathbf{H}_{k,l}\mathbf{a}_{k,n,l} + \mathbf{v}_n = \mathbf{H}\mathbf{a}_n + \mathbf{v}_n, \tag{14}$$

where $\mathbf{a}_{k,n,l} = \mathbf{s}_{k,n,l} - (1/M)\mathbf{1}_M$. For shorter notation, we denote the information symbol in $\mathbf{s}_{k,n,l}$ simply by $I$ after ignoring its time and user dependence. It takes values $0,\ldots,M-1$ with equal probability $1/M$. Then

$$\mathbf{a}_{k,n,l} = [\delta(I-0),\ldots,\delta(I-(M-1))]^{\mathrm{T}} - \frac{1}{M}\mathbf{1}_M^{\mathrm{T}}. \tag{15}$$

To obtain the covariance of $\mathbf{z}_n$, it is necessary to find that of $\mathbf{a}_{k,n,l}$. We denote it by $\mathbf{A} = E\{\mathbf{a}_{k,n,l}\mathbf{a}_{k,n,l}^{\mathrm{T}}\}$. According to the distribution of $I$, it can be found that

$$\mathbf{A} = \frac{1}{M}\sum_{i=1}^{M}\left(\mathbf{e}_{M,i} - \frac{1}{M}\mathbf{1}_M\right)\left(\mathbf{e}_{M,i} - \frac{1}{M}\mathbf{1}_M\right)^{\mathrm{T}}. \tag{16}$$

After simplification, it becomes

$$\mathbf{A} = \frac{1}{M}\left(\mathbf{I}_M - \frac{1}{M}\mathbf{1}_M\mathbf{1}_M^{\mathrm{T}}\right) \tag{17}$$

which is easily shown to have rank $M-1$ since $(1/\sqrt{M})\mathbf{1}_M$ is a unitary vector. Its $(m_1,m_2)$th element is defined as $a_{m_1,m_2}$. The ideal covariance of $\mathbf{z}_n$ is then derived to follow

$$\mathbf{R} = E\{\mathbf{z}_n\mathbf{z}_n^{\mathrm{H}}\} = \sum_{k,l}\mathbf{H}_{k,l}\mathbf{A}\mathbf{H}_{k,l}^{\mathrm{H}} + \sigma_v^2\mathbf{I}_\nu. \tag{18}$$

After defining a rank-one matrix $\mathbf{G}_k = \mathbf{g}_k\mathbf{g}_k^{\mathrm{H}}$, it becomes

$$\mathbf{R} = \sum_{k,l,m_1,m_2} a_{m_1,m_2}\mathbf{C}_{k,l}\mathbf{T}_{m_1}\mathbf{G}_k\mathbf{T}_{m_2}^{\mathrm{H}}\mathbf{C}_{k,l}^{\mathrm{H}} + \sigma_v^2\mathbf{I}_\nu. \tag{19}$$

As in [19], a vectored form is convenient to handle in the CM context. Define

$$\mathbf{r} = \mathrm{vec}(\mathbf{R}), \qquad \mathbf{x}_k = \mathrm{vec}(\mathbf{G}_k), \qquad \mathbf{x} = [\mathbf{x}_1^{\mathrm{T}},\ldots,\mathbf{x}_K^{\mathrm{T}},\sigma_v^2]^{\mathrm{T}}. \tag{20}$$

Using the property of vec [25], we obtain

$$\begin{aligned} \mathbf{r} &= \mathbf{S}\mathbf{x}, \\ \mathbf{S} &= [\mathbf{S}_1,\ldots,\mathbf{S}_K,\mathrm{vec}\,(\mathbf{I}_\nu)], \\ \mathbf{S}_k &= \sum_{l,m_1,m_2} a_{m_1,m_2}(\mathbf{C}_{k,l}\mathbf{T}_{m_2})^* \otimes (\mathbf{C}_{k,l}\mathbf{T}_{m_1}). \end{aligned} \tag{21}$$

Therefore, $\mathbf{r}$ can be matched with its estimate $\hat{\mathbf{r}}$ from $N$ data vectors in a vector form:

$$\hat{\mathbf{x}} = \arg\min ||\mathbf{r} - \hat{\mathbf{r}}||^2, \qquad \hat{\mathbf{r}} = \mathrm{vec}\,(\hat{\mathbf{R}}), \tag{22}$$

$$\hat{\mathbf{R}} = \frac{1}{N}\sum_{n=1}^{N}(\mathbf{y}_n - \hat{\bar{\mathbf{y}}})(\mathbf{y}_n - \hat{\bar{\mathbf{y}}})^{\mathrm{H}}. \tag{23}$$

Considering (21), the solution to (22) is given by

$$\hat{\mathbf{x}} = \mathbf{Q}\hat{\mathbf{r}}, \quad \mathbf{Q} = (\mathbf{S}^{\mathrm{H}}\mathbf{S})^{-1}\mathbf{S}^{\mathrm{H}}. \tag{24}$$

Once $\mathbf{x}$ is estimated, $\mathbf{x}_k$ can be extracted. Then $\mathbf{G}_k$ is reconstructed by the reverse vec operation. These operations can be described by

$$\begin{aligned} \hat{\mathbf{G}}_k &= [(\mathbf{e}_{q,1}^{\mathrm{T}} \otimes \mathbf{I}_q)\hat{\mathbf{x}}_k,\ldots,(\mathbf{e}_{q,q}^{\mathrm{T}} \otimes \mathbf{I}_q)\hat{\mathbf{x}}_k], \\ \hat{\mathbf{x}}_k &= [\mathbf{e}_{K,k}^{\mathrm{T}} \otimes \mathbf{I}_{q^2}, \mathbf{0}_{q^2\times 1}]\hat{\mathbf{x}}. \end{aligned} \tag{25}$$

Using (24), we can relate $\hat{\mathbf{G}}_k$ to $\hat{\mathbf{r}}$ as follows:

$$\hat{\mathbf{G}}_k = [\mathbf{A}_{k,1}\hat{\mathbf{r}},\ldots,\mathbf{A}_{k,q}\hat{\mathbf{r}}], \tag{26}$$

where

$$\mathbf{A}_{k,i} = (\mathbf{e}_{q,i}^{\mathrm{T}} \otimes \mathbf{I}_q)[\mathbf{e}_{K,k}^{\mathrm{T}} \otimes \mathbf{I}_{q^2}, \mathbf{0}_{q^2\times 1}]_{q^2\times(q^2K+1)}\mathbf{Q} \tag{27}$$

for $i = 1,\ldots,q$. Once $\hat{\mathbf{G}}_k$ is obtained, channel vector $\mathbf{g}_k$ can be estimated from its singular value decomposition (SVD) by finding the singular vector corresponding to its maximum singular value. That singular vector becomes an estimate of $\mathbf{g}_k$ up to a multiplicative scalar.

### 3.3. Complexity

From either (13) or (24), it is observed that all users' channel vectors can be estimated simultaneously. Although either matrix $\mathbf{W}$ or matrix $\mathbf{Q}$ can be precomputed offline, it still incurs possibly high complexity in multiplication. In the LS method, matrix $\mathcal{C}$ bas dimensionality $\nu\times qK$. Thus matrix $\mathbf{W}$ has dimensionality $qK\times\nu$. Multiplication of $\mathbf{W}$ by a vector of length $\nu$ in (13) requires complexity $O(\nu qK)$. In the CM method, matrix $\mathbf{S}$ has dimensionality $\nu^2\times(q^2K+1)$. Then matrix $\mathbf{Q}$ has dimensionality $(q^2K+1)\times\nu^2$. Therefore, (24) requires $(q^2K+1)\nu^2$ multiplications. The SVD on a $q\times q$ matrix $\hat{\mathbf{G}}_k$ incurs complexity $O(q^3)$. Total complexity to obtain a channel estimate by the CM method is about $O(\nu^2q^2K+q^3)$. For a long channel with large $q$, this complexity is mainly dominated by the SVD operation. In a case when only one user is of interest at a time, such as building an MMSE receiver to detect one user's signal, channel estimation complexity can be reduced by partitioning matrix $\mathbf{W}$ or $\mathbf{Q}$

row wise and taking entries corresponding to that user only. For example, consider estimation of channel vector of user 1 by the CM method. Extract the first $q^2$ rows of $\mathbf{Q}$ which is then multiplied by $\hat{\mathbf{r}}$ to obtain $\hat{\mathbf{x}}_1$. In such a way, computation load is reduced by about $K-1$ times.

Computational complexity can be further reduced by applying some adaptive techniques. For example in both methods, a stochastic gradient technique [26] can be easily applied based on corresponding cost functions (12) and (22). To reduce complexity from SVD of batch processing in order to obtain the singular vector of rank-one matrix $\mathbf{G}_k$ corresponding to its maximum singular value, power method has been shown to be very efficient [27, 28]. Subspace tracking technique [29] is also an excellent candidate for complexity reduction in implementation. Study of different adaptive implementations is not the focus of the current paper, but constitutes an interesting future research topic.

## 4. CHANNEL ESTIMATION PERFORMANCE

In this section, we study channel identifiability conditions. We also derive covariance and mean square error (MSE) of a channel estimate from each of the proposed channel estimators when the mean and covariance of the received data are not perfectly known but estimated by (11) and (23), respectively, from $N$ noisy data vectors.

### 4.1. Channel identifiability

Our solution (13) or (24) exists only when the corresponding matrix is invertible. Therefore, the LS channel estimator is unique if matrix $\mathcal{C}$ has full column rank. According to (10), it depends on TH codes of all users. It can only be checked for any assigned set of hopping sequences. Similarly, the CM channel estimator is unique if matrix $\mathbf{S}$ has full column rank. According to (21), it depends on time-hopping codes of all users as well [19]. It is also interesting to compare these two methods in terms of the maximum number of channels (users) that can be estimated. Matrix $\mathcal{C}$ has dimension $\nu \times Kq$. Then $K \le \lfloor \nu/q \rfloor$. However, $\mathbf{S}$ has size $\nu^2 \times (Kq^2+1)$. Correspondingly, $K$ approximately satisfies $K \le \lfloor (\nu/q)^2 \rfloor$. Usually, $\nu \gg q$. Hence, the CM method allows more simultaneous users than the LS method regarding channel estimation, but it has higher complexity since it uses SOS of the zero-mean data rather than easily estimated first-order statistic in the LS method. When symbol detection is of interest, a sufficient condition for all inputs to be detected is that a channel matrix which contains all columns of $\mathbf{H}_{k,l}$ (for all $k,l$) has full column rank. It thus imposes one more condition on $K$.

### 4.2. Performance of the LS estimator

It is observed that the LS channel estimation error is due to the error in estimating the mean of data vector in (11). From (13), we obtain

$$\delta\mathbf{g} = \mathbf{W}\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{y}_n - \bar{\mathbf{y}}\right) = \frac{1}{N}\mathbf{W}\sum_{n=1}^{N}\mathbf{z}_n. \tag{28}$$

Then covariance of $\delta\mathbf{g}$ becomes

$$\mathrm{COV}(\delta\mathbf{g}) = \frac{1}{N^2}\mathbf{W}\left(\sum_{n_1,n_2} E\{\mathbf{z}_{n_1}\mathbf{z}_{n_2}^{\mathrm{H}}\}\right)\mathbf{W}^{\mathrm{H}}. \tag{29}$$

Assume $\mathbf{z}_n$ constitutes a sequence of independent vectors. Then due to its zero mean, (29) is simplified as

$$\mathrm{COV}(\delta\mathbf{g}) = \frac{1}{N}\mathbf{W}\mathbf{R}\mathbf{W}^{\mathrm{H}}. \tag{30}$$

The covariance of $\delta\mathbf{g}_k$ is the $k$th diagonal subblock of this matrix and the MSE is the trace of the corresponding matrix.

### 4.3. Performance of the CM estimator

The CM channel estimation error is due to an estimation error in the data covariance $\mathbf{R}$ in (23) or equivalently $\mathbf{r}$. From our definition of $\mathbf{G}_k$, $\mathbf{g}_k$ is an eigenvector corresponding to its unique nonzero eigenvalue. If $\hat{\mathbf{r}}$ has an estimation error $\delta\mathbf{r} = \hat{\mathbf{r}} - \mathbf{r}$ due to finite $N$, then an error is introduced to $\hat{\mathbf{G}}_k$, and finally to our channel estimator. We derive the channel MSE as an explicit function of $N$ and system parameters next.

From (26), $\mathbf{G}_k$ is perturbed by $\delta\mathbf{G}_k$ as

$$\delta\mathbf{G}_k = [\mathbf{A}_{k,1}\delta\mathbf{r}, \ldots, \mathbf{A}_{k,q}\delta\mathbf{r}]. \tag{31}$$

Then the first-order perturbation in its eigenvector $\mathbf{g}_k$ becomes [30]

$$\delta\mathbf{g}_k \approx \mathbf{\Pi}_{\mathbf{g}_k}^{\perp}\delta\mathbf{G}_k\mathbf{g}_k, \qquad \mathbf{\Pi}_{\mathbf{g}_k}^{\perp} = \mathbf{\Sigma}_k\mathbf{\Sigma}_k^{\mathrm{H}}, \tag{32}$$

where $\mathbf{\Sigma}_k$ is in size of $q \times (q-1)$ and spans a $(q-1)$-dimensional subspace orthogonal to $\mathbf{g}_k$. Substituting (31) into (32), we obtain

$$\delta\mathbf{g}_k \approx \mathbf{\Gamma}_k\delta\mathbf{r}, \qquad \mathbf{\Gamma}_k = \mathbf{\Pi}_{\mathbf{g}_k}^{\perp}\sum_{i=1}^{q} g_k(i)\mathbf{A}_{k,i}. \tag{33}$$

Then the autocovariance of channel estimate becomes

$$\mathrm{COV}(\delta\mathbf{g}_k) = E\{\delta\mathbf{g}_k\delta\mathbf{g}_k^{\mathrm{H}}\} \approx \mathbf{\Gamma}_k\mathbf{\Phi}(\delta\mathbf{r})\mathbf{\Gamma}_k^{\mathrm{H}}, \tag{34}$$

where $\mathbf{\Phi}(\delta\mathbf{r}) = E\{\delta\mathbf{r}\delta\mathbf{r}^{\mathrm{H}}\}$ is the covariance of $\delta\mathbf{r}$. It depends on data model (7) and covariance estimation method in (23). According to [31, equation (12)], applying multilinearity and additivity properties of cumulant [32], and properties of $\otimes$ [25], the following proposition can be proved. Noticing that $M$ inputs in $\mathbf{a}_{k,n,l}$ are all real and dependent due to the same information, the last term in [31, equation (12)] does not diminish with real inputs.

Proposition 1. *If the channel model follows* (7) *and the data covariance is estimated from* $N$ *independent data vectors as* (23)*, then for a real system (all quantities are real),*

$$\begin{aligned}\mathbf{\Phi}(\delta\mathbf{r}) = {} & \frac{(N-1)^2}{N^3}\mathbf{K}_z + \frac{N-1}{N^2}\mathbf{R}\otimes\mathbf{R} \\ & + \frac{(N-1)^2}{N^3}\mathbf{B}_{1r}\odot\mathbf{B}_{1r}^{\mathrm{T}} \\ & + \frac{1}{N^2}\mathbf{r}\mathbf{r}^{\mathrm{T}} + \frac{N-1}{N^3}\mathbf{B}_{2r}\end{aligned} \tag{35}$$

*while for complex channel and noise,*

$$\begin{aligned}\mathbf{\Phi}(\delta\mathbf{r}) = {} & \frac{(N-1)^2}{N^3}\mathbf{K}_z + \frac{N-1}{N^2}\mathbf{R}^*\otimes\mathbf{R} \\ & + \frac{(N-1)^2}{N^3}\mathbf{B}_{1c}\odot\mathbf{B}_{1c}^{\mathrm{H}} \\ & + \frac{1}{N^2}\mathbf{r}\mathbf{r}^{\mathrm{H}} + \frac{N-1}{N^3}\mathbf{B}_{2c},\end{aligned} \tag{36}$$

*where*

$$\mathbf{B}_{1r} = (\mathbf{I}_\nu\otimes\mathbf{1}_\nu)\mathbf{R}(\mathbf{1}_\nu^{\mathrm{T}}\otimes\mathbf{I}_\nu), \tag{37}$$

$$\mathbf{B}_{2r} = \mathbf{R}\diamond\mathbf{R}, \tag{38}$$

$$\mathbf{B}_{1c} = (\mathbf{I}_\nu\otimes\mathbf{1}_\nu)\sum_{k,l}(\mathbf{H}_{k,l}^*\mathbf{A}\mathbf{H}_{k,l}^{\mathrm{H}})(\mathbf{1}_\nu^{\mathrm{T}}\otimes\mathbf{I}_\nu), \tag{39}$$

$$\mathbf{B}_{2c} = (\mathbf{H}^*\otimes\mathbf{H})(\mathcal{A}\diamond\mathcal{A})(\mathbf{H}^*\otimes\mathbf{H})^{\mathrm{H}},\quad \mathcal{A} = \mathbf{I}_L\otimes\mathbf{A}, \tag{40}$$

$\mathbf{K}_z$ *is the fourth-order cumulant matrix of* $\mathbf{z}_n$ *and is related to the fourth-order cumulant matrix* $\mathbf{K}_a$ *of input vector* $\mathbf{a}_{k,n,l}$ *as follows:*

$$\mathbf{K}_z = \sum_{k,l}(\mathbf{H}_{k,l}^*\otimes\mathbf{H}_{k,l})\mathbf{K}_a(\mathbf{H}_{k,l}^*\otimes\mathbf{H}_{k,l})^{\mathrm{H}}, \tag{41}$$

$$\begin{aligned}\mathbf{K}_a = {} & \frac{1}{M}\sum_{i=1}^{M}(\tilde{\mathbf{e}}_{M,i}\tilde{\mathbf{e}}_{M,i}^{\mathrm{T}})\otimes(\tilde{\mathbf{e}}_{M,i}\tilde{\mathbf{e}}_{M,i}^{\mathrm{T}}) \\ & - \mathrm{vec}(\mathbf{A})\,\mathrm{vec}(\mathbf{A})^{\mathrm{T}} - \mathbf{A}\otimes\mathbf{A} - \mathbf{B}_3\odot\mathbf{B}_3^{\mathrm{T}},\end{aligned} \tag{42}$$

*where* $\tilde{\mathbf{e}}_{M,i} = \mathbf{e}_{M,i} - (1/M)\mathbf{1}_M$ *and* $\mathbf{B}_3 = (\mathbf{I}_M\otimes\mathbf{1}_M)\mathbf{A}(\mathbf{1}_M^{\mathrm{T}}\otimes\mathbf{I}_M)$.

For the proof of the proposition, see the appendix.

*Remarks* 1. The statistical results in Proposition 1 are different from those in a CDMA system [19] since the current data model (14) shows a completely different structure although similarities exist. $M$ virtual inputs in $\mathbf{a}_{k,n,l}$ for each user in each symbol period are from the same information symbol. Meanwhile, all entries in $\mathbf{a}_{k,n,l}$ are real with only one positive dominant element. Overall, $\mathbf{\Phi}(\delta\mathbf{r})$ by either (35) or (36) is in the order of $1/N$ since the last two terms in each equation is in the order of $1/N^2$ and can be omitted for large $N$. Except the first term that depends on the fourth-order statistics of the data vector, all other terms are related to the SOS.

As expected, performance of the LS estimator depends only on the SOS of the data vector as shown by (30) because the estimator uses the first-order statistics. It is also in the order of $1/N$. Notice that there is no approximation in derivation of that result. The CM estimator however depends on the fourth-order statistics since the estimator is based on the covariance of data. Approximation under large sample-size ($N$) assumption has been made in order to obtain perturbation of eigenvectors of a matrix. All previous analyses are under an assumption of independent data vectors. It can be realized by taking one data vector every two symbol periods (only for channel estimation purpose not for detection purpose), or one data vector every symbol period but with removal of ISI. If successive (dependent) data vectors are adopted, similar analysis is still performable in principle. But corresponding results will become more complicated than before due to nonzero cross-correlation between any two successive data vectors contributed by ISI. The current analytical results provide a guideline for evaluation of estimators' performance. High agreement between simulation results with dependent data vectors and current analytical results (under data independence assumption) is still observed.

## 5. SIMULATION

In our simulation, the system adopts binary PPM modulation with $N_c = 4$ and $N_f = 10$. The monocycle pulse is chosen as a normalized second derivative of the Gaussian pulse with pulse duration $D_g = 0.7$ nanosecond, as in [7]. Modulation delay parameter is set to be $\sigma = D_g$. Unless otherwise specified, four users with equal transmitting power are active in the system. Their TH sequences are randomly generated. Multipath channels have time-delay resolution of $D_g$ with the maximum delay spread to be one frame duration. Channel gains for different users are modeled as independent Gaussian random variables and weighted by linearly decreasing weights [16]. All channel gains at the pulse rate are estimated. The corresponding normalized channel MSE for user $k$ defined as $\|\mathbf{g}_k - \hat{\mathbf{g}}_k\|^2/\|\mathbf{g}_k\|^2$ is used to measure the performance of each channel estimator.

Averaged channel estimation errors over 100 independent realizations for one set of channel parameters of one user are illustrated in Figure 1. Either successive (with ISI) or independent (ISI free) data vectors can be collected. As expected, MSE decreases monotonically with the increase of data length $N$. When $N$ is above 300, both the LS and CM estimators achieve a normalized MSE less than $10^{-2}$. Under the current channel condition, the CM estimator is consistently better than the LS estimator for all $N$. Analytical curves are also plotted and show high consistency with their experimental counterparts (with independent vectors). Data length effect on bit error rate (BER) is assessed in Figure 2 at $E_b/N_0 = 15$ dB using the subspace-based MMSE receiver [33]. Although BER decreases continuously when $N$ increases, the improvement becomes marginal when $N$ exceeds 400.

In Figure 3, we plot MSE as a function of SNR with data length $N = 1000$. For fixed $N$, the CM estimator outperforms the LS estimator at high SNR while the situation is reversed at low SNR. When the noise power is very low, both curves show an MSE floor due to finite data length $N$. Based on estimated channel coefficients, we examine the BER in Figure 4 for each of RAKE, ZF, conventional MMSE (based on direct

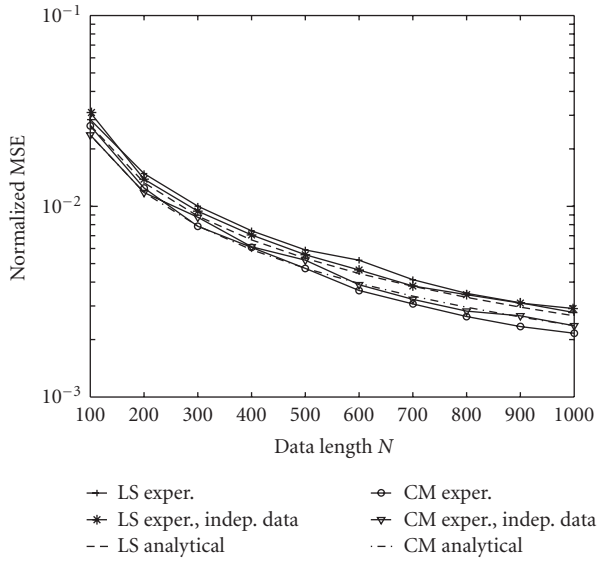



Figure 1: Channel estimation error under 15 dB SNR.

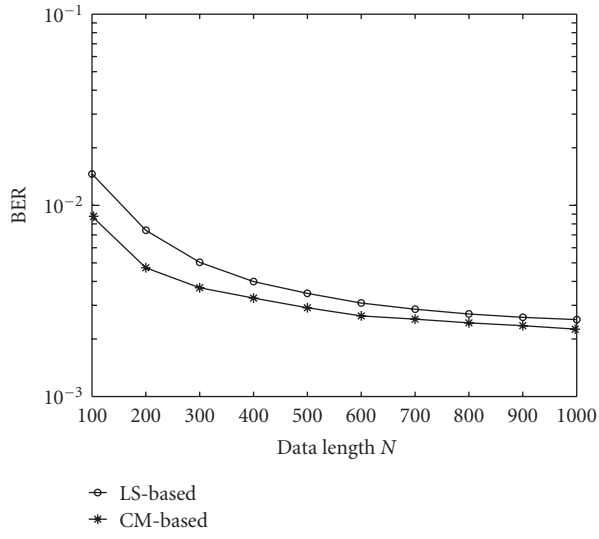



Figure 2: Data length effect on MMSE receiver's performance.

inversion of data covariance matrix), and subspace MMSE receivers. For comparison, corresponding ideal receivers assuming true channel parameters and correlation matrix, if applicable, are included. Among the four linear receivers, the ZF and subspace MMSE receivers show similar performance and outperform the others. RAKE and ZF receivers can achieve performance close to their ideal ones, indicating sufficient accuracy in our channel estimates. But for the conventional MMSE receiver, a clear gap is observed between experimental and ideal ones at high SNR mainly due to amplification of noise in practical conditions. However, its subspace variant almost achieves its ideal performance by removing the noise effect. Moreover, receivers based on the CM channel estimator are slightly better than those based on the LS estimator at high SNR.

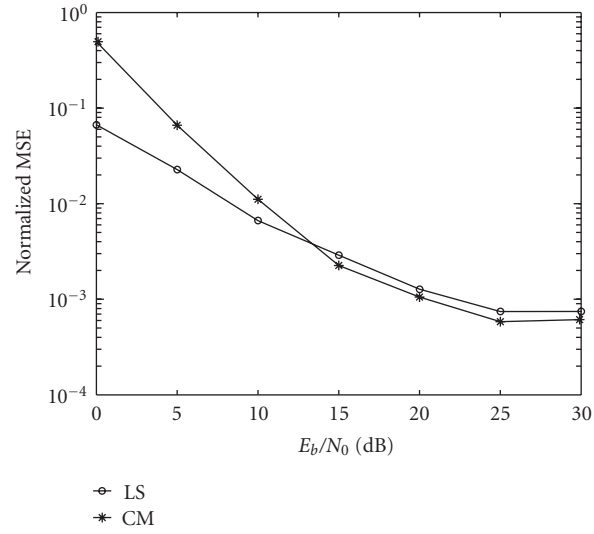



Figure 3: MSE versus SNR.

Figures 5 and 6 illustrate the impact of the number of users on channel estimation and symbol detection. Both MSE and BER degrade when more users become active in the system. But the degradation is gradual. Comparing two methods, the CM method degrades more slowly than the LS method. Therefore, the CM estimator is more robust in the sense that it can support more users. Indeed, when the number of users is more than 10, the channel cannot be estimated using the LS method because the estimator in (13) does not exist.

We have assumed equal power for all users in the above simulations. In the scenarios where stronger interferers exist in the system or the desired transmitter is far away from the receiver, equal transmitting power is not realistic. Hence, we examine the near-far effect in Figures 7 and 8. All interfering users except the desired user have the same power. Near-far ratio (NFR) is defined as the ratio of the power of each interfering user to that of the desired user. We can see that the MSE performance degrades little for the LS estimator, and slowly for the CM estimator when the NFR increases. BER almost keeps unchanged with change of the NFR. This shows that our proposed methods are strongly resistant to the near-far effect.

In the last experiment, we compare our proposed channel estimation method with the ML data-aided (ML-DA) and non-data-aided (ML-NDA) methods in [17]. First, multipath channels for all synchronized users are generated with path delays $3lD_g$ ($l = 1, 2, 3$). Path gains are assigned similarly as [17]. Unlike [17], to consider delay estimation error and gain estimation error separately, we assess estimation performance using channel MSE defined before and plot it with respect to SNR in Figure 9. We consider a single-user system in Figure 9a and a five-user system in Figure 9b. It is interesting to note the MSE of each ML method does not change with SNR. Both proposed LS method and CM method outperform the ML methods when the SNR is above 7 dB.

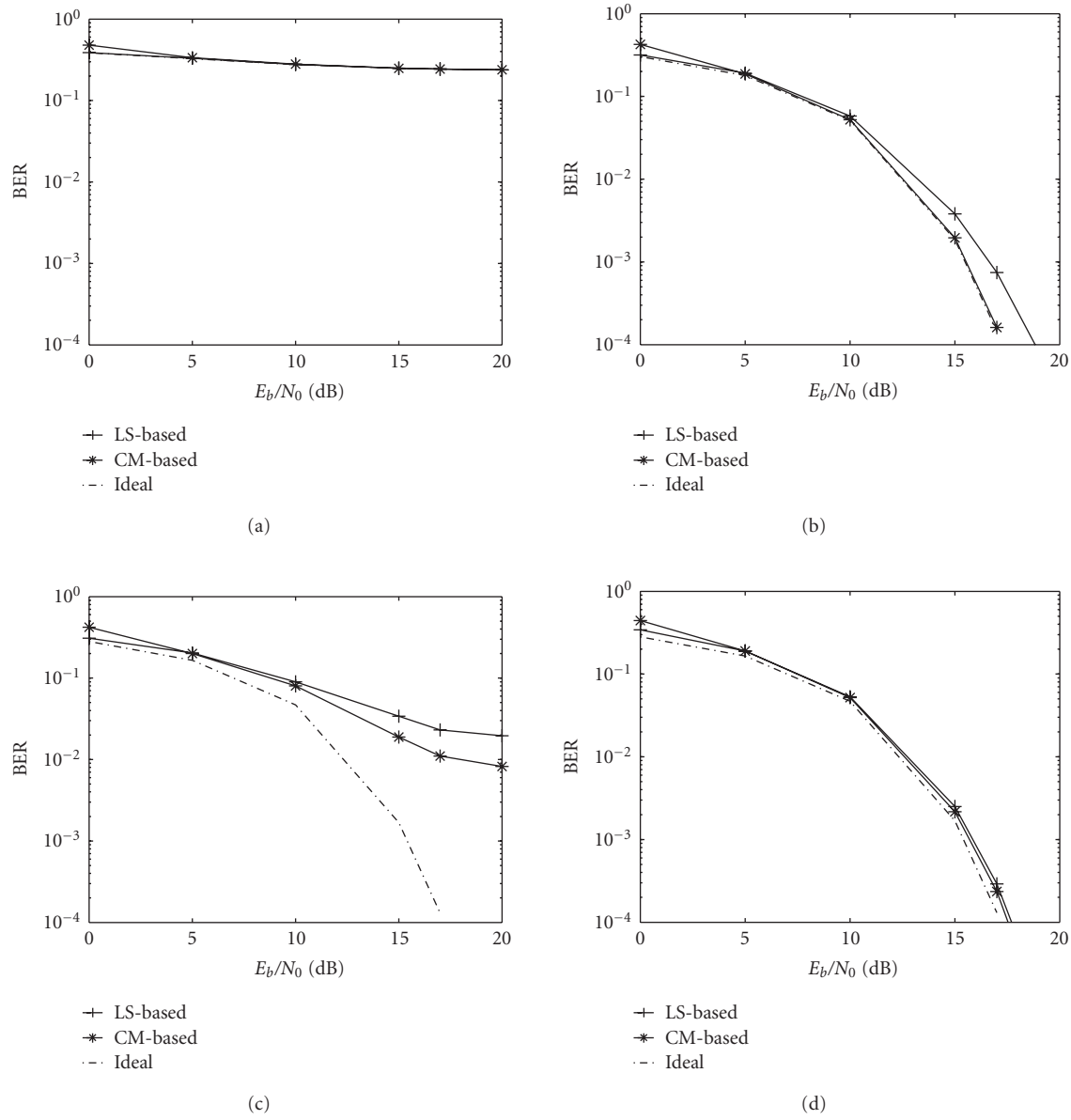



Figure 4: Detection performance of different receivers: (a) RAKE receiver, (b) ZF receiver, (c) conventional MMSE receiver, and (d) subspace MMSE receiver.

The ML methods are very sensitive to the number of users because of multiuser interference. Its estimation performance degrades dramatically in the multiuser system while the LS method and CM method degrade moderately, leading to significant performance gaps between them. Even in a single-user system, the ML methods do not always perform better mainly due to self-interference such as intersymbol interference and interframe interference. To compare the BER performance, two types of receivers are considered: three-finger RAKE receiver and subspace-based MMSE receiver. Results for a single-user system and a multiuser system are presented in Figures 10a and 10b, respectively. Despite the minor estimation difference existing in a single-user system among various methods, each kind of receiver (RAKE or MMSE) based on different estimation results shows almost the same performance. However, in an MA environment, advantages of our methods become obvious. Receivers based on the LS or CM method result in much lower BER than those based on ML methods. The MMSE receiver is always superior to the RAKE receiver. To see how asynchronism affects performance of different methods, we assume the first path delay of each user is unknown which is randomly and independently generated over half-frame duration. Thus each channel vector becomes longer due to overparameterization by some leading zeros. Compared with a synchronous (known delay) case, the number of unknowns

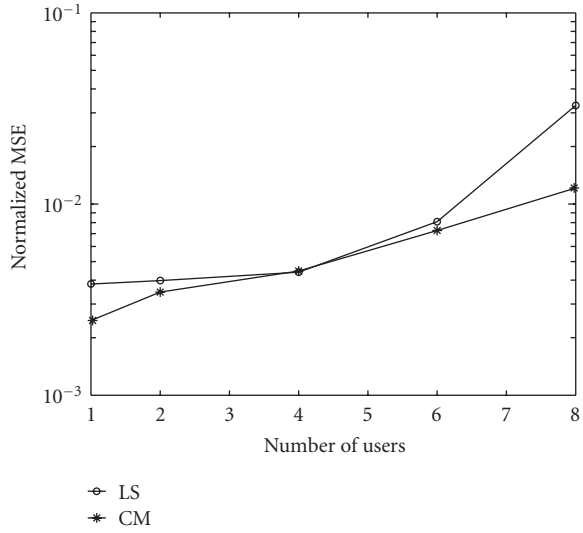



Figure 5: Number of users' effect on channel MSE.

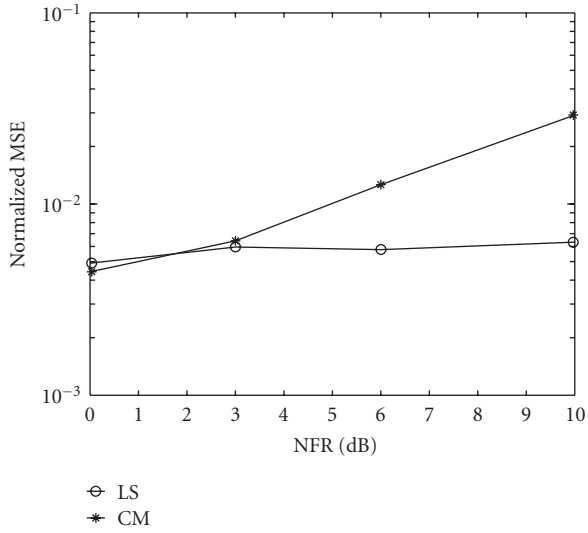



Figure 7: MSE versus NFR.

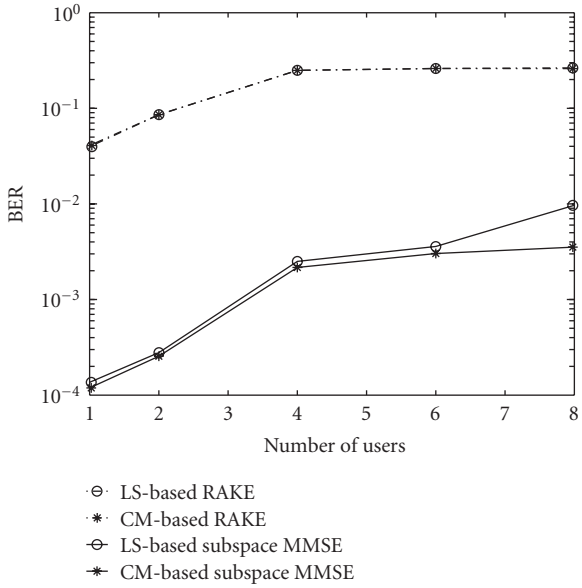



Figure 6: Number of users' effect on detection performance of the RAKE receiver and the MMSE receiver.

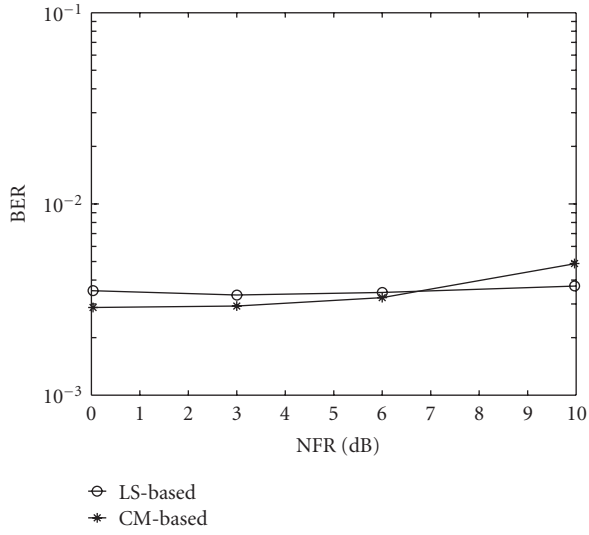



Figure 8: BER of MMSE receivers versus NFR.

to be estimated for each user increases. Correspondingly, we consider a four-user system in order to satisfy the identifiability condition for the LS method. To minimize redundancy, only channel MSE results for different SNRs using different $N$ are presented in Figure 11. It is observed that the proposed methods degrade compared with the synchronous case for the same $N$. They require larger sample size to achieve similar performance as before. The CM method performs best at high SNR. Meanwhile, both proposed methods still outperform the ML methods at high SNR for each examined $N$, although ML methods improve due to randomization of interference contributed by random propagation delays.

## 6. CONCLUSION

Based on a pulse-rate discrete-time UWB signal model, we propose LS and CM methods to blindly estimate channels in a multipath and MA environment by exploring first-order statistic or SOS of the received data vector. Multiuser receivers can then be constructed with acquired channel information. Channel estimation performance measured by MSE is analyzed and verified by simulations. Meanwhile, detection performance of corresponding multiuser receivers is numerically evaluated. The influence of some important factors, including data length, SNR, number of users, and NFR, on the system performance is investigated as well. The CM method performs better at high SNR and can support more users. However, the LS method has lower complexity and is found

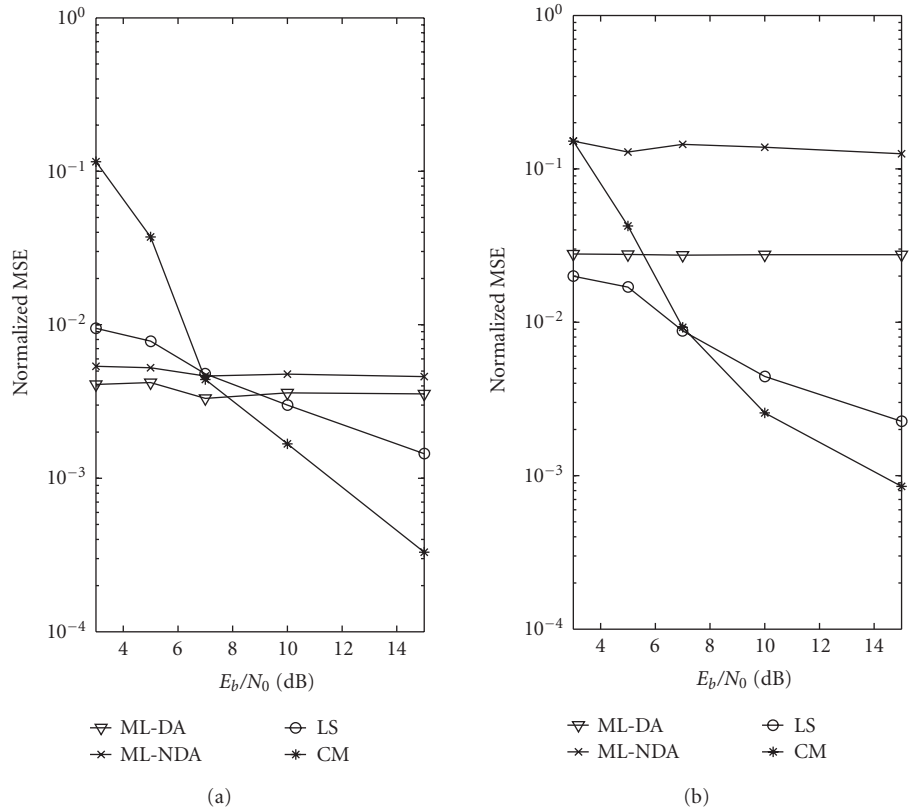



Figure 9: Channel estimation error of different methods using $N = 600$. (a) Single-user system. (b) Multiuser system.

to be more resistant to the near-far effect. Both methods outperform the existing ML channel estimation methods in a multiple access environment when the sample size is large.[1]

## APPENDIX

## PROOF OF PROPOSITION 1

The covariance $\mathbf{\Phi}(\delta\mathbf{r})$ depends on $\delta\mathbf{r} = \hat{\mathbf{r}} - \mathbf{r}$ which in turn depends on $\hat{\mathbf{R}}$ by (23) since $\hat{\mathbf{r}} = \text{vec}(\hat{\mathbf{R}})$. We thus first relate $\hat{\mathbf{R}}$ to $\mathbf{y}_n$ and then $\mathbf{z}_n$, which shows explicit dependence of system parameters.

Substituting the mean estimate (11) into (23), we obtain

$$\hat{\mathbf{R}} = \frac{1}{N}\sum_n \mathbf{y}_n\mathbf{y}_n^{\mathrm{H}} - \frac{1}{N^2}\sum_{n_1,n_2}\mathbf{y}_{n_1}\mathbf{y}_{n_2}^{\mathrm{H}}. \tag{A.1}$$

After substituting $\mathbf{y}_n$ by $\mathbf{z}_n + \bar{\mathbf{y}}$ according to (11), (A.1) can be simplified as

$$\hat{\mathbf{R}} = \frac{1}{N}\sum_n \mathbf{z}_n\mathbf{z}_n^{\mathrm{H}} - \frac{1}{N^2}\sum_{n_1,n_2}\mathbf{z}_{n_1}\mathbf{z}_{n_2}^{\mathrm{H}}. \tag{A.2}$$

In fact, (A.2) is the same as a typical covariance estimate

$$\hat{\mathbf{R}} = \frac{1}{N}\sum_n (\mathbf{z}_n - \hat{\bar{\mathbf{z}}})(\mathbf{z}_n - \hat{\bar{\mathbf{z}}})^{\mathrm{H}}, \quad \hat{\bar{\mathbf{z}}} = \frac{1}{N}\sum_n \mathbf{z}_n, \tag{A.3}$$

although we estimate $\mathbf{R}$ directly from $\mathbf{y}_n$ as (23). For convenience, define $\mathbf{r}_n = \text{vec}(\mathbf{z}_n\mathbf{z}_n^{\mathrm{H}})$. From (A.2), we have

$$\hat{\mathbf{r}} = \frac{N-1}{N^2}\sum_n \mathbf{r}_n - \frac{1}{N^2}\sum_{n_1\neq n_2}\text{vec}\left(\mathbf{z}_{n_1}\mathbf{z}_{n_2}^{\mathrm{H}}\right). \tag{A.4}$$

Due to zero mean and independence assumption on $\mathbf{z}_n$ at different $n$, the mean of $\hat{\mathbf{r}}$ is found to be $E\{\hat{\mathbf{r}}\} = (1 - 1/N)\mathbf{r}$. Then $\mathbf{\Phi}(\delta\mathbf{r})$ can be expanded into

$$\mathbf{\Phi}(\delta\mathbf{r}) = E\{(\hat{\mathbf{r}} - \mathbf{r})(\hat{\mathbf{r}} - \mathbf{r})^{\mathrm{H}}\} = E\{\hat{\mathbf{r}}\hat{\mathbf{r}}^{\mathrm{H}}\} - \left(1 - \frac{2}{N}\right)\mathbf{r}\mathbf{r}^{\mathrm{H}}. \tag{A.5}$$

It thus suffices to derive $E\{\hat{\mathbf{r}}\hat{\mathbf{r}}^{\mathrm{H}}\}$ for further simplification of $\mathbf{\Phi}(\delta\mathbf{r})$.

If we substitute (A.4), express $\text{vec}(\mathbf{z}_{n_1}\mathbf{z}_{n_2}^{\mathrm{H}})$ by $\mathbf{z}_{n_2}^* \otimes \mathbf{z}_{n_1}$, and invoke the property of $\otimes$ [25], we obtain

$$\begin{aligned} E\{\hat{\mathbf{r}}\hat{\mathbf{r}}^{\mathrm{H}}\} = {} & \left(\frac{N-1}{N^2}\right)^2 \sum_{n_1,n_2} E\{\mathbf{r}_{n_1}\mathbf{r}_{n_2}^{\mathrm{H}}\} \\ & + \frac{1}{N^4}\sum_{\substack{n_1\neq n_2 \\ n_3\neq n_4}} E\{(\mathbf{z}_{n_2}\mathbf{z}_{n_4}^{\mathrm{H}})^* \otimes (\mathbf{z}_{n_1}\mathbf{z}_{n_3}^{\mathrm{H}})\}. \end{aligned} \tag{A.6}$$

[1] The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the US Government.

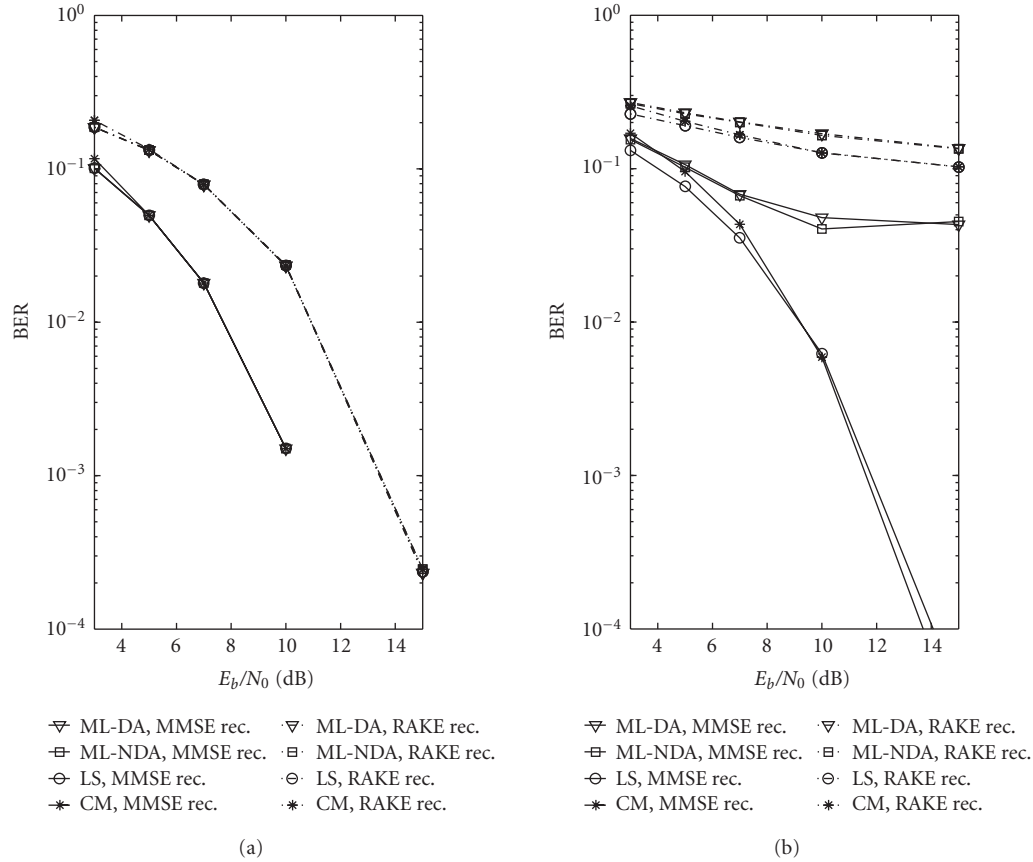



Figure 10: Detection performance based on different channel estimation results. (a) Single-user system. (b) Multiuser system.

The term $\sum_{n_1,n_2} E\{\mathbf{r}_{n_1}\mathbf{r}_{n_2}^{\mathrm{H}}\}$ becomes $NE\{\mathbf{r}_n\mathbf{r}_n^{\mathrm{H}}\}+(N^2-N)\mathbf{r}\mathbf{r}^{\mathrm{H}}$. In the second term, there are only two different cases which give nonzero contributions because of the zero mean of $\mathbf{z}_n$ and independence assumption: $n_2 = n_4$, $n_1 = n_3$ but $n_1 \neq n_2$; $n_2 = n_3$, $n_1 = n_4$ but $n_1 \neq n_2$. They correspondingly yield

$$(N^2-N)\mathbf{R}^* \otimes \mathbf{R} + (N^2-N)E\{(\mathbf{z}_{n_2}^*\mathbf{z}_{n_1}^{\mathrm{T}}) \otimes (\mathbf{z}_{n_1}\mathbf{z}_{n_2}^{\mathrm{H}})\} \quad \text{(A.7)}$$

where $n_1 \neq n_2$. Therefore, (A.6) becomes

$$\begin{aligned} E\{\hat{\hat{\mathbf{r}}}\mathbf{r}^{\mathrm{H}}\} = {} & \frac{(N-1)^2}{N^3}E\{\mathbf{r}_n\mathbf{r}_n^{\mathrm{H}}\} - \frac{(N-1)^3}{N^3}\mathbf{r}\mathbf{r}^{\mathrm{H}} \\ & + \frac{N-1}{N^3}\mathbf{R}^* \otimes \mathbf{R} + \frac{N-1}{N^3}\mathbf{B}_2, \end{aligned} \quad \text{(A.8)}$$

where for notational convenience, we may replace $\mathbf{z}_{n_1}$ and $\mathbf{z}_{n_2}$ by $\mathbf{u}$ and $\mathbf{w}$, respectively, and correspondingly define

$$\mathbf{B}_2 = E\{(\mathbf{u}^*\mathbf{w}^{\mathrm{T}}) \otimes (\mathbf{w}\mathbf{u}^{\mathrm{H}})\}. \quad \text{(A.9)}$$

In (A.8), both $E\{\mathbf{r}_n\mathbf{r}_n^{\mathrm{H}}\}$ and $\mathbf{B}_2$ need further simplification. We begin with $\mathbf{B}_2$ next.

In (A.9), vectors $\mathbf{u}$ and $\mathbf{w}$ both have zero mean and covariance $\mathbf{R}$. They are independent because $n_1 \neq n_2$. If $\mathbf{u}$ and $\mathbf{w}$ are real, then we denote $\mathbf{B}_2$ by $\mathbf{B}_{2r}$:

$$\begin{aligned} \mathbf{B}_{2r} &= E\{(\mathbf{u}\mathbf{w}^{\mathrm{T}}) \otimes (\mathbf{w}\mathbf{u}^{\mathrm{T}})\} \\ &= E\{(\mathbf{I}_\nu\mathbf{u}\mathbf{w}^{\mathrm{T}}) \otimes (\mathbf{w}\mathbf{u}^{\mathrm{T}}\mathbf{I}_\nu)\} \\ &= E\{(\mathbf{I}_\nu \otimes \mathbf{w})(\mathbf{u} \otimes 1)(1 \otimes \mathbf{u}^{\mathrm{T}})(\mathbf{w}^{\mathrm{T}} \otimes \mathbf{I}_\nu)\} \\ &= E\{(\mathbf{I}_\nu \otimes \mathbf{w})(\mathbf{R} \otimes 1)(\mathbf{w}^{\mathrm{T}} \otimes \mathbf{I}_\nu)\} \\ &= E\{(\mathbf{R} \otimes \mathbf{w})(\mathbf{w}^{\mathrm{T}} \otimes \mathbf{I}_\nu)\}. \end{aligned} \quad \text{(A.10)}$$

After expressing $\mathbf{w}^{\mathrm{T}} \otimes \mathbf{I}_\nu$ by $[w_1\mathbf{I}_\nu, \ldots, w_\nu\mathbf{I}_\nu]$ based on elements of $\mathbf{w}$, (A.10) is simplified as

$$\mathbf{B}_{2r} = \mathbf{R} \diamond \mathbf{R} \quad \text{(A.11)}$$

which is (38). If $\mathbf{u}$ and $\mathbf{w}$ are complex, for example due to complex channels and noise, then we denote $\mathbf{B}_2$ by $\mathbf{B}_{2c}$. Under an assumption of circular symmetry of Gaussian noise and using data model (14) for $\mathbf{u}$ and $\mathbf{w}$, those terms involving noise disappear. Therefore, only one term in $\mathbf{B}_{2c}$

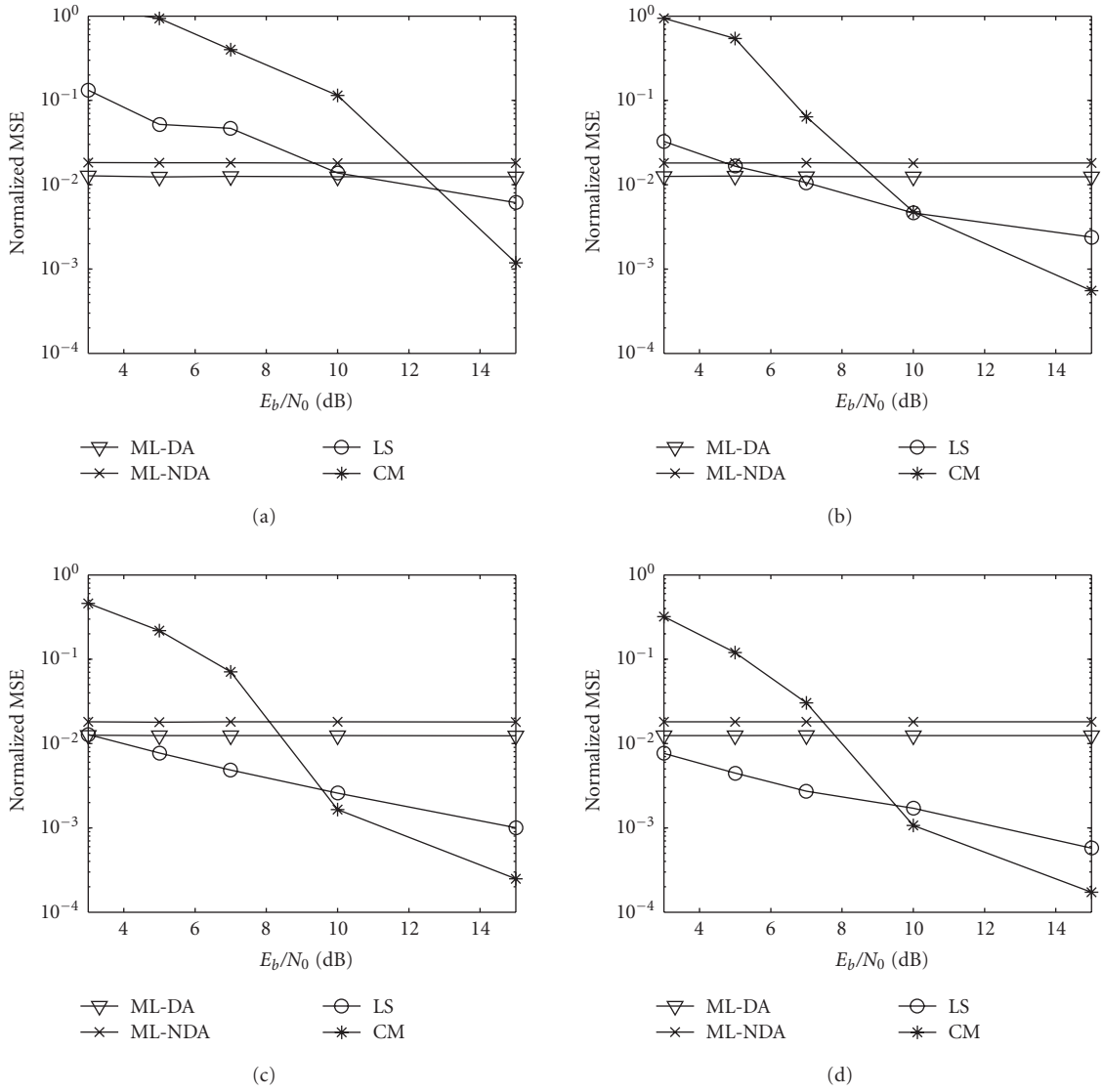



Figure 11: Comparison of different channel estimators in an asynchronous system: (a) $N = 600$, (b) $N = 1500$, (c) $N = 3000$, and (d) $N = 5000$.

is left as

$$\begin{aligned}\mathbf{B}_{2c} &= E\{(\mathbf{H}^* \mathbf{a}_{n_1} \mathbf{a}_{n_2}^{\mathrm{T}} \mathbf{H}^{\mathrm{T}}) \otimes (\mathbf{H} \mathbf{a}_{n_2} \mathbf{a}_{n_1}^{\mathrm{T}} \mathbf{H}^{\mathrm{H}})\} \\ &= (\mathbf{H}^* \otimes \mathbf{H}) E\{(\mathbf{a}_{n_1} \mathbf{a}_{n_2}^{\mathrm{T}}) \otimes (\mathbf{a}_{n_2} \mathbf{a}_{n_1}^{\mathrm{T}})\} (\mathbf{H}^{\mathrm{T}} \otimes \mathbf{H}^{\mathrm{H}}) \\ &= (\mathbf{H}^* \otimes \mathbf{H})(\mathcal{A} \diamond \mathcal{A})(\mathbf{H}^* \otimes \mathbf{H})^{\mathrm{H}},\end{aligned} \tag{A.12}$$

where we have applied (A.10) and (A.11) to real vectors $\mathbf{a}_{n_1}$ and $\mathbf{a}_{n_2}$ whose autocovariance is $\mathcal{A} = \mathbf{I}_L \otimes \mathbf{A}$. Then (40) follows.

We turn our attention to $E\{\mathbf{r}_n \mathbf{r}_n^{\mathrm{H}}\}$. Replacing $\mathbf{r}_n$ by $\mathbf{z}_n^* \otimes \mathbf{z}_n$ and applying [31, equation (12)], we obtain

$$E\{\mathbf{r}_n \mathbf{r}_n^{\mathrm{H}}\} = \mathbf{K}_z + \mathbf{r}\mathbf{r}^{\mathrm{H}} + \mathbf{R}^* \otimes \mathbf{R} + \mathbf{B}_1 \odot \mathbf{B}_1^{\mathrm{H}}, \tag{A.13}$$

where $\mathbf{K}_z$ is the fourth-order cumulant of $\mathbf{z}_n$ and

$$\mathbf{B}_1 = E\{(\mathbf{z}_n^* \otimes \mathbf{1}_\nu)(\mathbf{1}_\nu \otimes \mathbf{z}_n)^{\mathrm{H}}\}. \tag{A.14}$$

Similar to simplification of $\mathbf{B}_2$, we can simplify $\mathbf{B}_1$ as follows:

$$\begin{aligned}\mathbf{B}_1 &= E\{(\mathbf{I}_\nu \mathbf{z}_n^* \otimes \mathbf{1}_\nu)(\mathbf{1}_\nu^{\mathrm{T}} \otimes \mathbf{z}_n^{\mathrm{H}} \mathbf{I}_\nu)\} \\ &= E\{(\mathbf{I}_\nu \otimes \mathbf{1}_\nu)(\mathbf{z}_n^* \otimes 1)(1 \otimes \mathbf{z}_n^{\mathrm{H}})(\mathbf{1}_\nu^{\mathrm{T}} \otimes \mathbf{I}_\nu)\} \\ &= (\mathbf{I}_\nu \otimes \mathbf{1}_\nu) E\{\mathbf{z}_n^* \mathbf{z}_n^{\mathrm{H}}\} (\mathbf{1}_\nu^{\mathrm{T}} \otimes \mathbf{I}_\nu).\end{aligned} \tag{A.15}$$

If $\mathbf{z}_n$ is real, then $E\{\mathbf{z}_n^* \mathbf{z}_n^{\mathrm{H}}\} = \mathbf{R}$ and (A.15) becomes (37). If $\mathbf{z}_n$ is complex, then $E\{\mathbf{z}_n^* \mathbf{z}_n^{\mathrm{H}}\} = \sum_{k,l} (\mathbf{H}_{k,l}^* \mathbf{A} \mathbf{H}_{k,l}^{\mathrm{H}})$ and (A.15) becomes (39).

The cumulant $\mathbf{K}_z$ is related to statistics of $\mathbf{a}_{k,n,l}$. According to data model (14), applying multilinearity and additivity properties of cumulant [32], and noticing the zero cumulant of Gaussian noise, we obtain (41), where $\mathbf{K}_a$ is the fourth-order cumulant of $\mathbf{a}_{k,n,l}$. $\mathbf{K}_a$ can be similarly found by applying [31, equation (12)] again to an $M \times 1$ real vector $\mathbf{a}_{k,n,l}$.

Since $\mathbf{a}_{k,n,l}$ has zero mean and covariance $\mathbf{A}$, the result becomes

$$\begin{aligned}\mathbf{K}_a = E\{(\mathbf{a}_{k,n,l}\otimes\mathbf{a}_{k,n,l})(\mathbf{a}_{k,n,l}\otimes\mathbf{a}_{k,n,l})^{\mathrm{T}}\}\\ -\operatorname{vec}(\mathbf{A})\operatorname{vec}(\mathbf{A})^{\mathrm{T}}-\mathbf{A}\otimes\mathbf{A}-\mathbf{B}_3\odot\mathbf{B}_3^{\mathrm{T}},\end{aligned}\tag{A.16}$$

where $\mathbf{B}_3$ has a similar form as $\mathbf{B}_1$ in (A.15):

$$\mathbf{B}_3 = E\{(\mathbf{a}_{k,n,l}\otimes\mathbf{1}_M)(\mathbf{1}_M\otimes\mathbf{a}_{k,n,l})^{\mathrm{T}}\}.\tag{A.17}$$

Changing $\nu$ in (A.15) to $M$, we obtain $\mathbf{B}_3 = (\mathbf{I}_M\otimes\mathbf{1}_M)\mathbf{A}(\mathbf{1}_M^{\mathrm{T}}\otimes\mathbf{I}_M)$. The expectation in (A.16) can be derived from the definition of $\mathbf{a}_{k,n,l}$ in (15). Since the information symbol $I$ takes values $0,\ldots,M-1$ with probability $1/M$, we have

$$\begin{aligned}&E\{(\mathbf{a}_{k,n,l}\otimes\mathbf{a}_{k,n,l})(\mathbf{a}_{k,n,l}\otimes\mathbf{a}_{k,n,l})^{\mathrm{T}}\}\\ &\quad=\frac{1}{M}\sum_{i=1}^{M}(\widetilde{\mathbf{e}}_{M,i}\widetilde{\mathbf{e}}_{M,i}^{\mathrm{T}})\otimes(\widetilde{\mathbf{e}}_{M,i}\widetilde{\mathbf{e}}_{M,i}^{\mathrm{T}}).\end{aligned}\tag{A.18}$$

Substituting (A.18) into (A.16), (42) follows. Applying the above results and substituting (A.8) into (A.5), we complete the proof.

## ACKNOWLEDGMENTS

This paper was prepared through collaborative participation in the Communications and Networks Consortium sponsored by the US Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011. The US Government is authorized to reproduce and distribute reprints for government purposes notwithstanding any copyright notation thereon. This paper was presented in part at the IEEE Topical Conference on Wireless Communication Technology, Hawaii, October 2003, and the IEEE UWBST Conference, Virginia, November 2003.

## REFERENCES

[1] G. F. Ross, *The transient analysis of multiple beam feed networks for array systems*, Ph.D. dissertation, Polytechnic Institute of Brooklyn, Brooklyn, NY, USA, 1963.

[2] G. F. Ross, "The transient analysis of certain TEM mode four-port networks," *IEEE Transactions on Microwave Theory and Techniques*, vol. 14, no. 11, pp. 528–542, 1966.

[3] H. F. Harmuth, *Transmission of Information by Orthogonal Functions*, Springer-Verlag, New York, NY, USA, 1969.

[4] G. F. Ross, "A time domain criterion for the design of wideband radiation elements," *IEEE Trans. Antennas and Propagation*, vol. 16, no. 3, pp. 355–356, 1968.

[5] C. L. Bennett and G. F. Ross, "Time-domain electromagnetics and its applications," *Proceedings of the IEEE*, vol. 66, no. 3, pp. 299–318, 1978.

[6] Federal Communications Commission (FCC), "Revision of part 15 of the commission's rules regarding ultra-wideband transmission systems," First Report and Order, ET Docket 98-153, FCC 02-48; Adopted:February 2002; Released:April 2002.

[7] M. Z. Win and R. A. Scholtz, "Impulse radio: how it works," *IEEE Communications Letters*, vol. 2, no. 2, pp. 36–38, 1998.

[8] R. Fontana, A. Ameti, E. Richley, L. Beard, and D. Guy, "Recent advances in ultra wideband communications systems," in *Proc. IEEE Conference on Ultra Wideband Systems and Technologies (UWBST '02)*, pp. 129–133, Baltimore, Md, USA, May 2002.

[9] R. A. Scholtz, "Multiple access with time-hopping impulse modulation," in *Proc. of the IEEE Military Communications Conference (MILCOM '93)*, vol. 2, pp. 447–450, Boston, Mass, USA, October 1993.

[10] M. Z. Win and R. A. Scholtz, "On the robustness of ultra-wide bandwidth signals in dense multipath environments," *IEEE Communications Letters*, vol. 2, no. 2, pp. 51–53, 1998.

[11] M. Z. Win and R. A. Scholtz, "On the energy capture of ultrawide bandwidth signals in dense multipath environments," *IEEE Communications Letters*, vol. 2, no. 9, pp. 245–247, 1998.

[12] M. Z. Win, "Spectral density of random UWB signals," *IEEE Communications Letters*, vol. 6, no. 12, pp. 526–528, 2002.

[13] M. Z. Win and R. A. Scholtz, "Characterization of ultra-wide bandwidth wireless indoor channels: a communication-theoretic view," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 9, pp. 1613–1627, 2002.

[14] D. Cassioli, M. Z. Win, and A. R. Molisch, "The ultra-wide bandwidth indoor channel: from statistical model to simulations," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1247–1257, 2002.

[15] C. J. Le Martret and G. B. Giannakis, "All-digital impulse radio with multiuser detection for wireless cellular systems," *IEEE Trans. Communications*, vol. 50, no. 9, pp. 1440–1450, 2002.

[16] L. Yang and G. B. Giannakis, "Multistage block-spreading for impulse radio multiple access through ISI channels," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 9, pp. 1767–1777, 2002.

[17] V. Lottici, A. D'Andrea, and U. Mengali, "Channel estimation for ultra-wideband communications," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 9, pp. 1638–1645, 2002.

[18] L. Yang and G. B. Giannakis, "Optimal pilot waveform assisted modulation for ultrawideband communications," *IEEE Transactions on Wireless Communications*, vol. 3, no. 4, pp. 1236–1249, 2004.

[19] Z. Xu, "Asymptotically near-optimal blind estimation of multipath CDMA channels," *IEEE Trans. Signal Processing*, vol. 49, no. 9, pp. 2003–2017, 2001.

[20] D. Porcino and W. Hirt, "Ultra-wideband radio technology: potential and challenges ahead," *IEEE Communications Magazine*, vol. 41, no. 7, pp. 66–74, 2003.

[21] Q. Li and L. Rusch, "Multiuser detection for DS-CDMA UWB in the home environment," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 9, pp. 1701–1711, 2002.

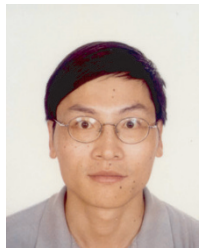
[22] J. R. Foerster, "The performance of a direct-sequence spread ultrawideband system in the presence of multipath, narrowband interference, and multiuser interference," in *Proc. IEEE Conference on Ultra Wideband Systems and Technologies (UWBST '02)*, pp. 87–91, Baltimore, Md, USA, May 2002.

[23] B. Sadler and A. Swami, "On the performance of UWB and DS-spread spectrum communication systems," in *Proc. IEEE Conference on Ultra Wideband Systems and Technologies (UWBST '02)*, pp. 289–292, Baltimore, Md, USA, May 2002.

[24] M. Hamalainen, V. Hovinen, R. Tesi, J. H. Iinatti, and M. Latva-aho, "On the UWB system coexistence with GSM900, UMTS/WCDMA, and GPS," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 9, pp. 1712–1721, 2002.

[25] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, Academic Press, San Diego, Calif, USA, 2nd edition, 1985.

[26] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, Upper Saddle River, NJ, USA, 3rd edition, 1996.
[27] G. Golub and C. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, Md, USA, 2nd edition, 1989.
[28] T. Moon and W. Stirling, *Mathematical Methods and Algorithms for Signal Processing*, Prentice Hall, Upper Saddle River, NJ, USA, 2000.
[29] B. Yang, "Projection approximation subspace tracking," *IEEE Trans. Signal Processing*, vol. 43, no. 1, pp. 95–107, 1995.
[30] Z. Xu, "Perturbation analysis for subspace decomposition with applications in subspace-based algorithms," *IEEE Trans. Signal Processing*, vol. 50, no. 11, pp. 2820–2830, 2002.
[31] A.-J. van der Veen, "Statistical performance analysis of the algebraic constant modulus algorithm," *IEEE Trans. Signal Processing*, vol. 50, no. 12, pp. 3083–3097, 2002.
[32] A. J. van der Veen, "Algebraic constant modulus algorithms," in *Signal Processing Advances in Wireless & Mobile Communications. Volume 2: Trends in Single- and Multi-User Systems*, G. B. Giannakis, Y. Hua, P. Stoica, and L. Tong, Eds., chapter 3, pp. 89–130, Prentice Hall, Upper Saddle River, NJ, USA, 2000.
[33] X. Wang and H. Poor, "Blind multiuser detection: a subspace approach," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 677–690, 1998.



**Zhengyuan Xu** received both the B.S. and M.S. degrees in electronic engineering from Tsinghua University, Beijing, China, in 1989 and 1991, respectively, and the Ph.D. degree in electrical engineering from Stevens Institute of Technology, Hoboken, NJ, USA, in 1999. From 1991 to 1996, he worked as an Engineer and a Department Manager at the Tsinghua Unisplendour Group Corp., Tsinghua University. Since 1999, he has been with the Department of Electrical Engineering, University of California, Riverside, as an Assistant Professor. His current research interests include detection and estimation theory, spread-spectrum and ultra-wideband wireless technology, multiuser communications, and ad hoc and wireless sensor networking. Dr. Xu received the Outstanding Student Award and the Motorola Scholarship from Tsinghua University, and the Peskin Award from Stevens Institute of Technology. He also received the Academic Senate Research Award and the Regents' Faculty Award from the University of California, Riverside. He has served as a Session Chair and a Technical Program Committee Member for international conferences. He is an IEEE Senior Member, a Member of the IEEE Signal Processing Society's Technical Committee on Signal Processing for Communications, and an Associate Editor for the IEEE Transactions on Vehicular Technology and the IEEE Communications Letters.

**Jin Tang** received the B.S. degree in electrical engineering and business administration from Beijing University of Aeronautics and Astronautics in 1995, and the M.S. degree in electrical engineering from the University of California, Riverside, in 2003. He is now a Ph.D. candidate in the Electrical Engineering Department, the University of California, Riverside. His research interests include channel estimation, and receiver design for ultra-wideband communication systems.

**Ping Liu** received the B.S. and M.S. degrees in electronic engineering from Sichuan University, Chendu, China, in 1990 and 1993, respectively, the M.E. degree in electrical and electronic engineering from Nanyang Technological University, Singapore, in 1999, and the Ph.D. degree in electrical engineering from the University of California, Riverside, in 2004. From April 1999 to August 1999, she worked as a Research Engineer at Kent Ridge Digital Labs, Singapore. Since August 2004, she has been an Assistant Professor in the Department of Electrical Engineering, Arkansas Tech University, Russellville, Ark. Her current research interests are in the general area of wireless communications, space-time coding, digital signal processing, and blind system identification.