Designing Training Sequences for Carrier Frequency Estimation in Frequency-Selective Channels

Young-Doo Kim, *Student Member, IEEE*, Jae Kun Lim, Changho Suh, *Member, IEEE*, and Yong H. Lee, *Senior Member, IEEE*

***Abstract*—A procedure for selecting a training sequence (TS) is developed for frequency estimation in frequency-selective channels. An expression for the unconditional Cramér–Rao bound (UCRB) is obtained by averaging the CRB for frequency estimation over the probability density function of Gaussian random channels. In addition, a necessary and sufficient condition for minimizing the UCRB is derived. Based on these results, a procedure for selecting a TS is developed. Through a computer search, binary TSs up to length 24 are found and tabulated. It is observed that periodic TSs tend to be selected when the TS length is twice the channel duration. Simulation results demonstrate that the proposed TSs can enhance the performance of the maximum likelihood (ML) frequency estimate.**

***Index Terms*—Frequency estimation, frequency-selective channels, training sequence (TS) design.**

I. Introduction

FREQUENCY OFFSET compensation and channel estimation are important functions of recent radio receivers. By recovering the carrier frequency at a receiver via signal processing, stringent requirements on the frequency stability of the transmitter and receiver oscillators can be relieved. The estimation of channel parameters is essential for systems with maximum likelihood (ML)-type detection. A typical data packet consists of user data and a training sequence (TS), which is used for synchronization and channel estimation. For such packets, much effort has been devoted to finding an optimal TS. In [1]–[6], TSs minimizing the variance of least squares (LS)-type and discrete Fourier transform (DFT)-based channel estimation are obtained. Also, using a lower bound on the capacity of training-based schemes enables the optimal placement and power allocation for the training symbols to be found [7], [8], and the length of the training interval for channel estimation to be optimized [9]. Optimal TSs for frequency estimation in frequency-selective channels are found in [10] and [11] based on the use of the Cramér–Rao Bound (CRB). Here, the optimal (minmax) condition for minimizing the maximum CRB resulting from the worst channel is derived [10], and a white sequence is shown to be optimal in the minmax sense for an asymptotic CRB [10], [11].

In this paper, we consider an alternative approach to selecting TSs for frequency estimation in frequency-selective channels. The proposed method is based on the use of an *unconditional* CRB (UCRB) derived by averaging the CRB for frequency estimation [12] over the probability density function (pdf) of Gaussian random channels. The conditions for minimizing the UCRB are thus derived, and a procedure is developed for finding TSs. The result is then applied to finding TSs for the ML frequency estimation. Through computer simulation, the advantage of the proposed TS is demonstrated over existing TSs, including white sequences.

The organization of this paper is as follows. Section II presents the system model. In Section III, a procedure for selecting TSs is developed based on the UCRB. Section IV outlines simulation results demonstrating the advantage of the proposed TSs. Finally, conclusions are drawn in Section V.

II. System Model

The baseband system model considered in this paper is shown in Fig. 1. Here, $a(j)$ denotes linearly modulated [phase shift keying (PSK) or quadrature amplitude modulation (QAM)] symbols; $g(t)$ is the baseband pulse shape; $w(t)$ is additive white Gaussian noise (AWGN); and Δf represents the carrier frequency offset. The output of the receiver filter sampled at $t = kT$ can be written as

$$r(k) = e^{j2\pi\nu k} \sum_{l=0}^{L-1} a(k-l)h_k(l) + w(k) \quad (1)$$

where $\nu = \Delta fT$ denotes the normalized frequency offset, and $h_k(l)$ is the impulse response of the equivalent channel at time k due to an impulse that is applied l time units earlier. As such, it describes $g(t), c(t)$, and the receiver filter in the discrete time domain, and its duration is L. $w(k)$ is assumed to be AWGN with a variance of σ_n^2.

Suppose that $N + L - 1$ training symbols $\{a(k)|k = -L + 1, \ldots, N-1\}$ are transmitted, and that $\{h_k(l)\}$ is fixed over the training period; i.e., $h_k(l) = h(l)$ for $k = -L+1, \ldots, N-1$. Ignoring the first $L-1$ samples, we define the received vector $\mathbf{r} = [r(0), r(1), \ldots, r(N-1)]^T$ that corresponds to the training sequence and channel vector $\mathbf{h} = [h(0), h(1), \ldots, h(L-1)]^T$. Then, (1) can be rewritten in vector form as

$$\mathbf{r} = \mathbf{\Gamma}(\nu)\mathbf{A}\mathbf{h} + \mathbf{w} \quad (2)$$

where $\mathbf{\Gamma}(\nu)$ is a diagonal matrix given by $\mathbf{\Gamma}(\nu) = \text{diag}[1, e^{j2\pi\nu}, \ldots, e^{j2\pi(N-1)\nu}]$; $\mathbf{A}$ is a N-by-L matrix with entries

Manuscript received November 3, 2003; revised October 10, 2004, May 22, 2005. This work was supported in part by the University Information Technology Research Center Program of the government of Korea. The review of this paper was coordinated by Dr. Y.-P. Eric Wang.

Y.-D. Kim and Y. H. Lee are with the Department of Electrical Engineering and Computer Science, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 305-701, Korea (e-mail: ckyd@stein.kaist.ac.kr; yohlee@ee.kaist.ac.kr).

J. K. Lim is with the Department of Computational Mathematics and Informatics, Hankyong National University, Kyonggido 456-749, Korea (e-mail: jaekun@hknu.ac.kr).

C. Suh is with Telecommunication R&D Center, Samsung Electronics Company, Ltd., Kyonggido 443-742, Korea (e-mail: becal.suh@samsung.com).

Digital Object Identifier 10.1109/TVT.2005.861173

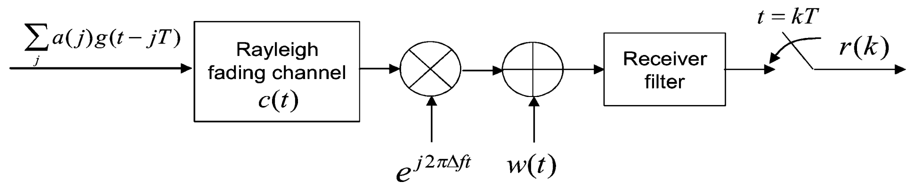


Fig. 1. Baseband system model.

$[\mathbf{A}]_{i,j} = a(i-j), 0 \le i \le N-1, 0 \le j \le L-1$; and $\mathbf{w} = [w(0), w(1), \ldots, w(N-1)]^T$. The training matrix $\mathbf{A}$ in (2) becomes a circulant matrix when $a(-L+j) = a(N-L+j)$ for $1 \le j \le L-1$. Throughout the paper, $\mathbf{A}$ is assumed to be circulant.

III. Training Sequence Design Based on UCRB

If we are considering the joint estimation of ν and $\mathbf{h}$, the CRB in the estimation of ν is given by [12], [13]

$$\mathrm{CRB} = \frac{\sigma_n^2}{2\mathbf{y}^H(\mathbf{I}_N - \mathbf{B})\mathbf{y}} \tag{3}$$

where $\mathbf{y} = 2\pi\mathbf{MAh}$, $\mathbf{M} = \mathrm{diag}\ [0, 1, \ldots, N-1]$, $\mathbf{B} = \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$, and $\mathbf{I}_N$ is the N-by-N identity matrix.[1] The UCRB is defined as UCRB $= \mathrm{E}_{\mathbf{h}}[\mathrm{CRB}]$, where $\mathrm{E}_{\mathbf{h}}[\cdot]$ denotes the expectation with respect to the *a priori* pdf of $\mathbf{h}$. It is shown in [13] that the UCRB provides a tighter bound than the *modified* CRB in [14].

Suppose that $\mathbf{h}$ is a zero-mean complex Gaussian random vector with the covariance matrix $\mathbf{C}_h = \mathrm{E}[\mathbf{hh}^H]$, which is assumed to be positive definite. The normalized channel vector $\mathbf{x}$ is given by $\mathbf{x} := \mathbf{C}_h^{-1/2}\mathbf{h}$, where $\mathbf{C}_h = \mathbf{C}_h^{1/2}\mathbf{C}_h^{1/2}$. Then $\mathbf{x}$ is a standard complex Gaussian random vector with a zero-mean, and $\mathrm{E}[\mathbf{xx}^H] = \mathbf{I}_L$. Using $\mathbf{h} = \mathbf{C}_h^{1/2}\mathbf{x}$, (3) can be rewritten as

$$\begin{aligned}\mathrm{CRB} &= \frac{\sigma_n^2/8\pi^2}{\mathbf{x}^H(\mathbf{C}_h^{1/2})^H\mathbf{A}^H\mathbf{M}(\mathbf{I}_N - \mathbf{B})\mathbf{MAC}_h^{1/2}\mathbf{x}} \\ &= \frac{\sigma_n^2/8\pi^2}{\mathbf{x}^H\mathbf{Px}}\end{aligned} \tag{4}$$

where $\mathbf{P}$ is an L-by-L matrix given by

$$\mathbf{P} = (\mathbf{C}_h^{1/2})^H\mathbf{A}^H\mathbf{M}(\mathbf{I}_N - \mathbf{B})\mathbf{MAC}_h^{1/2} \tag{5}$$

which is positive definite, because $(\mathbf{I}_N - \mathbf{B})^H(\mathbf{I}_N - \mathbf{B}) = (\mathbf{I}_N - \mathbf{B})$. To evaluate the UCRB, the CRB in (4) is averaged over the pdf of $\mathbf{x}$, as shown in the following lemma.

Lemma 1: Let $\mathbf{x} = [x_1, x_2, \ldots, x_L]^T$ and $\|\mathbf{x}\|^2 = \sum_{i=1}^L |x_i|^2 = \sum_{i=1}^L(|x_{i,1}|^2 + |x_{i,2}|^2)$, where $x_{i,1}$ and $x_{i,2}$ are the real and imaginary parts of x_i, respectively. Then

$$\mathrm{UCRB} = \frac{\sigma_n^2}{8\pi^2}f(\underline{\lambda}) \tag{6a}$$

where

$$f(\underline{\lambda}) = \int_{R^{2L}} \frac{\Phi(\mathbf{x})}{\lambda_1|x_1|^2 + \cdots + \lambda_L|x_L|^2}\,d\mathbf{x} \tag{6b}$$

and $\underline{\lambda} = [\lambda_1, \lambda_2, \ldots, \lambda_L]^T$; λ_is are eigenvalues of $\mathbf{P}$; $d\mathbf{x} = dx_{1,1}\ldots dx_{L,1}dx_{1,2}\ldots dx_{L,2}$; R denotes the real line; and $\Phi(\mathbf{x})$ is the joint pdf of $\mathbf{x}$, given by $\Phi(\mathbf{x}) = (2\pi)^{-L}\exp(-\|\mathbf{x}\|^2/2)$.

The proof of this lemma is presented in Appendix A. From the UCRB expression in (6), the following conjecture can be made.

Conjecture 1: The UCRB can be minimized by increasing each eigenvalue λ_i while maintaining

$$\lambda_1 = \lambda_2 = \cdots = \lambda_L. \tag{7}$$

The equality condition in (7) comes from the symmetry of $f(\underline{\lambda})$ in (6b) for all permutations of λ_i. The validity of this conjecture is examined through Lemma 2 and Theorem 1, as follows:

Lemma 2: Let Ω_m be a set of $\underline{\lambda}$ given by $\Omega_m = \{\underline{\lambda} \in (0,\infty)^L \,|\, \sum_{i=1}^L \lambda_i = m\}$ for some positive integer m. Then $f(\underline{\lambda})$ in (6) is strictly convex over Ω_m.

This lemma is proved in Appendix B. Due to Lemma 2, there exists a unique optimum, say $\underline{\lambda}_0$, that minimizes $f(\underline{\lambda})$ over Ω_m. The optimum vector $\underline{\lambda}_0$ and corresponding $\mathbf{P}$ matrix are found in the following theorem.

Theorem 1: Let $\mathbf{S}(m)$ be the set of all L-by-L positive definite matrices whose trace is equal to m, and $\mathbf{P} \in \mathbf{S}(m)$. Suppose that $L > 1$. Then, among the matrices in $\mathbf{S}(m)$, $\mathbf{P}$ yields the smallest UCRB if and only if $\mathbf{P} = (m/L)\mathbf{I}_L$ (or equivalently $\underline{\lambda}_0 = (m/L)[1, 1, \ldots, 1]^T$).

Proof: If $\underline{\lambda}$ is in Ω_m, then the trace of $\mathbf{P}$ is fixed by m. As such, it is sufficient to show that $f(\underline{\lambda})$ has a unique minimum at $\underline{\lambda}_o = m/L[1, 1, \ldots, 1]^T$ in the clearly convex set Ω_m. This is because when $\underline{\lambda} = m/L[1, 1, \ldots, 1]^T$, the eigenvalues of $\mathbf{Q}$ (and hence $\mathbf{P}$) are all m/L, and thus $\mathbf{P}$ becomes $(m/L)\mathbf{I}_L$. Let $\underline{\lambda}^{(i)}$ be the ith cyclic shift of an arbitrary point $\underline{\lambda} \in \Omega_m$. Then, $\underline{\lambda}^{(i)} \in \Omega_m$ for each i. Since $(m/L)[1, 1, \ldots, 1]^T = \sum_{i=0}^{L-1}(1/L)\underline{\lambda}^{(i)}$, then the convexity of f implies that

$$\begin{aligned}f\left(\frac{m}{L}[1,\ldots,1]^T\right) &= f\left(\sum_{i=0}^{L-1}\frac{1}{L}\underline{\lambda}^{(i)}\right) \le \sum_{i=0}^{L-1}\frac{1}{L}f(\underline{\lambda}^{(i)}) \\ &= \sum_{i=0}^{L-1}\frac{1}{L}f(\underline{\lambda}) = f(\underline{\lambda})\end{aligned}$$

where the second equality comes from the fact that $f(\underline{\lambda})$ is symmetric for all permutations of the coordinates of $\underline{\lambda}$. This completes the proof. □

Theorem 1 indicates that Conjecture 1 is indeed valid: to minimize the UCRB, each λ_i is maximized by maximizing the trace m $(m = \sum_{i=1}^L \lambda_i)$ while maintaining $\lambda_1 = \lambda_2 = \cdots = \lambda_L$.

[1] The CRB in (3) is referred to as the joint estimation CRB (JCRB) in [13].

Although Theorem 1 provides conditions for minimizing the UCRB, it is difficult to use this theorem in practice because $\mathbf{P}$ in (5) can hardly be a scaled identity matrix ($\mathbf{P} = (m/L)\mathbf{I}_L$)—it often happens that the optimal TSs do not exist, as shown in the following examples:

Example 1: Let $N = 8, L = 2$, and $\mathbf{C}_h^{-1/2} = \sqrt{2}\mathbf{I}_2$. Assume that TSs consist of quaternary-PSK (QPSK) symbols. Then there are 4^8 possible TSs. Through a computer search, it was found that 768 TSs maximized m and among them, 16 TSs yielded $\mathbf{P} = (42/2)\mathbf{I}_L$. One of the optimal TS is $\{a(k)|k = 0, 1, \ldots, 7\} = \{-1, j, 1, j, 1, j, -1, j\}$.

Example 2: Let $N = 8, L = 2$, and $\mathbf{C}_h^{-1/2} = \mathrm{diag}\,[1.1696, 1.9283]$. Assuming QPSK symbols, 768 TSs maximized m; however, none of them resulted in the desired $\mathbf{P}$ matrix ($\mathbf{P} \neq (m/L)\mathbf{I}_L$).

Example 3: Let $N = 16, L = 4$, and $\mathbf{C}_h^{-1/2} = \sqrt{4}\mathbf{I}_4$. Assuming binary-PSK (BPSK) symbols, four TSs maximized m; however, none of them resulted in the desired $\mathbf{P}$ matrix ($\mathbf{P} \neq (m/L)\mathbf{I}_L$).

Examples 2 and 3 illustrate the cases where optimal TSs do not exist. For such cases, an alternative approach is developed as follows:

Lemma 3: Let $\mathbf{P} \in \mathbf{S}(m)$ and $\det(\mathbf{P})$ and $tr(\mathbf{P})$ denote the determinant and trace of $\mathbf{P}$, respectively. Then, $\det(\mathbf{P})$ is the maximum of the determinants of the matrices in $\mathbf{S}(m)$ if and only if $\mathbf{P} = (m/L)\mathbf{I}_L$.

Proof:

$$\det(\mathbf{P}) \leq \prod_{i=1}^{L} [\mathbf{P}]_{i,i} \leq \left(\frac{1}{L}\sum_{i=1}^{L}[\mathbf{P}]_{i,i}\right)^L = \left(\frac{1}{L}tr(\mathbf{P})\right)^L = \left(\frac{m}{L}\right)^L$$

where the first inequality is the Hadamard inequality [15, p. 477] and the second inequality is the arithmetic-geometric mean inequality [15, p. 535]. Therefore, $\det(\mathbf{P}) \leq (m/L)^L$ with equality for $\mathbf{P} = (m/L)\mathbf{I}_L$. □

Now, Theorem 1 can be rewritten as follows:

Corollary 1: Among the matrices in $\mathbf{S}(m), \mathbf{P}$ yields the smallest UCRB if and only if $\det(\mathbf{P})$ is the maximum of the determinants associated with $\mathbf{S}(m)$.

This corollary suggests the following process for TS selection.

Process 1: Identify $\mathbf{S}(m_{\max})$, where $m_{\max}$ is the maximum of the traces of all possible $\mathbf{P}$ matrices in (5). Then, among $\mathbf{P}$ matrices in $\mathbf{S}(m_{\max})$, find the one with the largest $\det(\mathbf{P})$ and the corresponding TS.

A TS can always be found by Process 1; in particular, an optimal TS is selected whenever such a sequence exists. Next, an *ad hoc* scheme that is a simplified version of Process 1 is presented.

Process 2: Select a TS such that $\det(\mathbf{P})$ is maximized.

This process cannot guarantee to select an optimal TS. However, Process 2 performs better than Process 1 when optimal TSs do not exist. This can be shown by applying Processes 1 and 2 to Examples 2 and 3, for which optimal TSs do not exist. Table I shows the resulting TSs, corresponding traces, determinants, and CRBs (The binary TSs are expressed in hexadecimal form). Processes 1 and 2 yielded an identical TS for the case of Example 2. However, for Example 3, Process 2 resulted in a considerably smaller CRB value than Process 1. This happened because maximizing the traces prior to maximizing the determinants caused the determinant of the resulting $\mathbf{P}$ matrix to be approximately equal to zero, indicating that the equality condition in (7) was severely violated. Some additional experiments, which are not reported here, also showed the superiority of Process 2 over Process 1. The performance of Process 1 is sensitive to the existence of optimal TSs, while that of Process 2 is not. Furthermore, Process 2 does not require knowledge of the channel covariance matrix $\mathbf{C}_h$. This is because $\det(\mathbf{P}) = \det(\mathbf{C}_h)\det(\mathbf{A}^H\mathbf{M}(\mathbf{I}_N - \mathbf{B})\mathbf{M}\mathbf{A})$, and maximizing $\det(\mathbf{P})$ is equal to maximizing $\det(\mathbf{A}^H\mathbf{M}(\mathbf{I}_N - \mathbf{B})\mathbf{M}\mathbf{A})$. From these results, the following method, which is identical to Process 2, is proposed for TS selection.

TABLE I
TSs FOUND BY PROCESSES 1 AND 2 FOR EXAMPLES 2 AND 3, AND CORRESPONDING TRACES, DETERMINANTS, AND CRBs. HERE, CRBs WERE OBTAINED THROUGH 10^4 SIMULATION RUNS AT $E_b/N_0 = 20$ dB, USING (4)

		TS	$tr(\mathbf{P})$	$det(\mathbf{P})$	CRB
Example 2	Process 1	{j, j, j, -j, -j, j, -j, -j}	42	347	6.5×10^{-6}
	Process 2	{j, j, j, -j, -j, j, -j, -j}	42	347	6.5×10^{-6}
Example 3	Process 1	FDFF	336	$\simeq 0$	1.92×10^{-6}
	Process 2	E36E	331	4.67×10^{7}	5.1×10^{-7}

Proposed Method: Select a TS such that

$$\det(\mathbf{A}^H\mathbf{M}(\mathbf{I}_N - \mathbf{B})\mathbf{M}\mathbf{A}) \tag{8}$$

is maximized.

In [10], a scheme that is similar to the proposed method is derived by a minmax approach. Specifically, it maximizes

$$\lambda_{\min}[\mathbf{A}^H\mathbf{M}(\mathbf{I}_N - \mathbf{B})\mathbf{M}\mathbf{A}] \tag{9}$$

where $\lambda_{\min}[X]$ is the minimum eigenvalue of X, and will be referred to as the minmax method. Since the determinant of a square matrix is equal to the product of its eigenvalues, (8) can be thought of as a modification of (9).

A computer search was performed to find TSs satisfying the proposed and minmax conditions for a given TS length N and channel duration L. Table II lists the binary TSs found up to length 24. The TSs are expressed in hexadecimal form. The proposed and minmax methods produced identical sequences when $N = 8$ and $(N, L) = (12, 4)$, yet yielded different TSs for larger values of N ($N = 16, 20$, and 24). When $L = N/2$, the TSs found by the proposed method were either periodic with period L, or composed of a sequence of length L followed by its complement (the latter will be referred to as a complementarily periodic TSs). These TSs are particularly useful for the ML frequency estimate, because the computational complexity for implementing the estimate can be reduced dramatically by employing such TSs.

TABLE II
TSs FOUND BY PROPOSED AND MINMAX METHODS, WHERE N AND L DENOTE TS LENGTH AND CHANNEL DURATION, RESPECTIVELY

	L = N/4		L = N/3		L = N/2	
	Proposed	Minmax	Proposed	Minmax	Proposed	Minmax
N = 8	E4	E4	-	-	EE[1]	EE[1]
N = 12	D1D	D9D	E6E	E6E	EC4[2]	DF7[1]
N = 16	E36E	EEEE[1]	-	-	F4F4[1]	EE91
N = 20	E2476	77BB1	-	-	EDC48[2]	1DB89[2]
N = 24	D02BC8	E42E42	F6DC48	DBB891	FACFAC[1]	DD3DD3[1]

[1] The TS is periodic with period L.

[2] The TS consists of a sequence of length L followed by its complement.

IV. SIMULATION RESULTS

Computer simulations were conducted to examine the influence of the TS selection on the CRB and performance of the ML frequency estimate in [12], given by

$$\hat{\nu} = \arg\max_{\tilde{\nu}} \left\{ -\rho(0) + 2\mathrm{Re}\left[\sum_{m=0}^{\beta N-1} \rho(m) e^{-j2\pi m \tilde{\nu}} \right] \right\} \quad (10)$$

where we have the equation shown at the bottom of the page, $B(k,l)$ is the (k,l)th entry of the N-by-N matrix $\mathbf{B}$ in (3), and β is a positive integer. This ML estimate requires an exhaustive search over the range $|\nu| \le 0.5$. For a periodic TS consisting of two identical sequences $(N = 2L)$, the ML estimate reduces to

$$\hat{\nu} = \frac{1}{\pi N} \arg\left\{ \sum_{k=N/2}^{N-1} r(k) r^*(k - N/2) \right\} \quad (11)$$

which is a blockwise differential decoding-based operator that does not require any computer search. For the complementarily periodic TSs in Table II, the ML estimate is identical to the right-hand side (RHS) of (11) with the exception that "$r(k)r^*(k-N/2)$" is replaced with "$-r(k)r^*(k-N/2)$." When the signal-to-noise ratio (SNR) is large, the mean square error (MSE) of $\hat{\nu}$ can be approximated by

$$\mathrm{E}[(\hat{\nu} - \nu)^2 | \mathbf{h}] \cong \frac{\sigma_n^2}{2\mathbf{y}^H(\mathbf{I}_N - \mathbf{B})\mathbf{y}}. \quad (12)$$

Note that the RHS of (12) is identical to the CRB in (3). This fact indicates that the TS designed by the proposed method tends to minimize the MSE of the ML estimate for high SNR values.

In the simulation, two types of channel models were considered: a channel composed of independent identically distributed (i.i.d.) complex Gaussian random variables with a zero mean and variance $1/L$, and a typical urban (TU) channel model of the Global System for Mobile Communications (GSM) system with six paths, which was considered in [12]. The matrix $\mathbf{C}_h^{-1/2}$ for the i.i.d. channel was $\sqrt{L}\mathbf{I}_L$. The GSM channel response was represented as

$$h(l) = \sum_{i=0}^{5} \xi_i g_T(lT_s - \tau_i - t_0) \quad (13)$$

where $g_T(t)$ is the impulse response of the raised-cosine filter with a rolloff of 0.5; $\{\xi_i\}$ and $\{\tau_i\}$ are the attenuation and path delays, respectively; and t_0 is the timing phase. All the parameters for generating $h(l)$ in (13) were equal to those in [12]. For this GSM channel, L can be set at eight. The parameter β in (10) was 16. Three kinds of TSs were used: the proposed

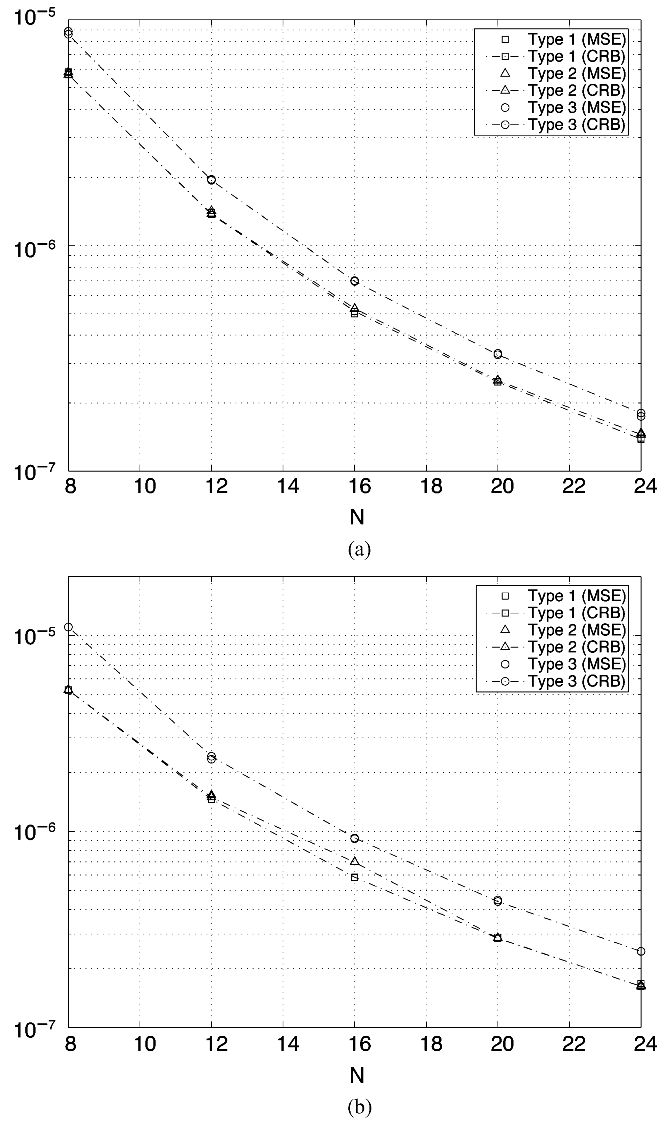
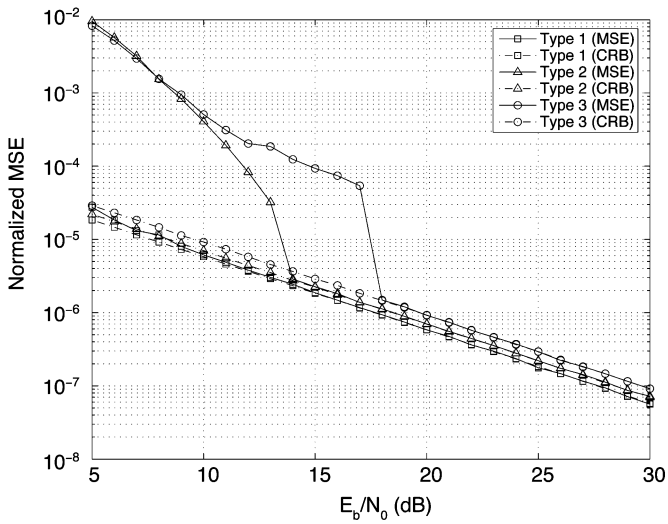
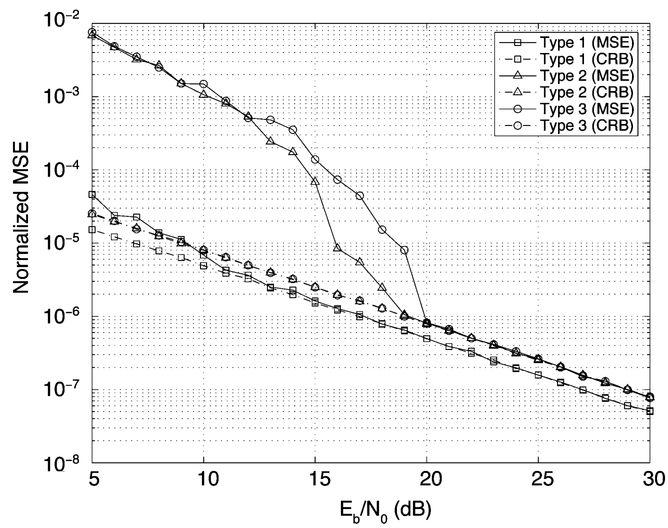


Fig. 2. Performance comparison ($E_b/N_0 = 20$ dB). (a) $L = N/4$. (b) $L = N/2$.

$$\rho(m) = \begin{cases} \sum_{k=m+1}^{N} B(k-m,k) r(k) r^*(k-m), & 0 \le m \le N-1 \\ 0, & \text{otherwise.} \end{cases}$$

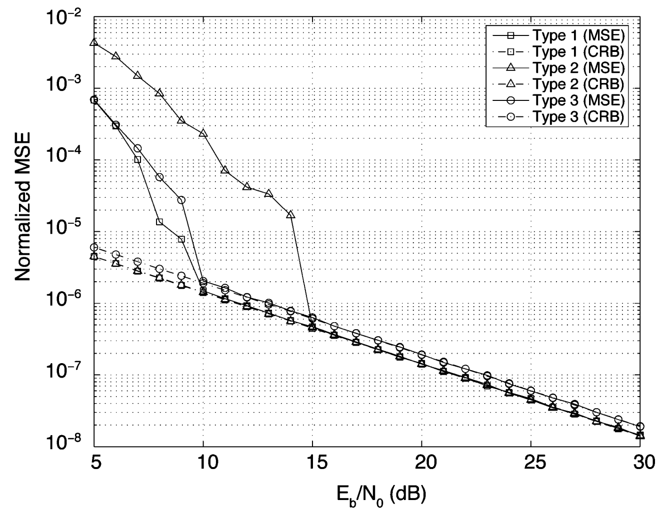


(a)

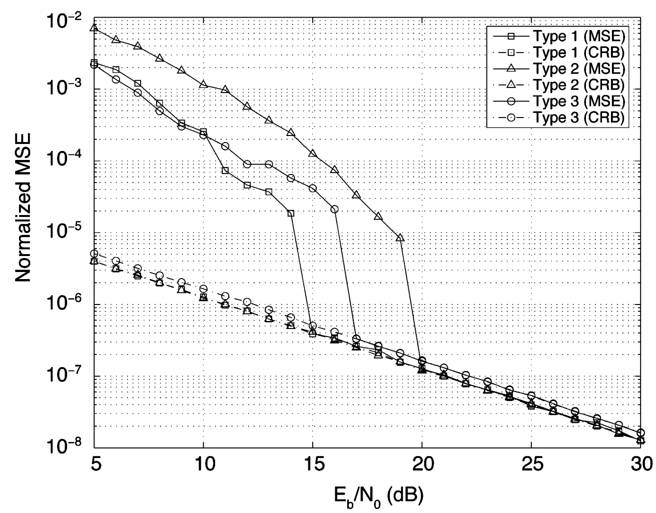


(b)

Fig. 3. Performance comparison $(N = 16, L = 8)$. (a) i.i.d. Gaussian channels. (b) GSM TU channels.



(a)



(b)

Fig. 4. Performance comparison $(N = 24, L = 8)$. (a) i.i.d. Gaussian channels. (b) GSM TU channels.

and minmax TSs in Table II, and a white TS.[2] These TSs will be referred to as Type 1 (proposed), Type 2 (minmax), and Type 3 (white) TSs. The CRB and MSE values were obtained through 10^4 trials. For each trial, different channels and a different white TS were generated, and the normalized frequency offset ν was taken randomly from the interval $[-0.04, 0.04]$.

Fig. 2 shows the CRBs and MSEs of $\hat{\nu}$ against N when the channel was i.i.d. Gaussian and the SNR per bit (E_b/N_0) was 20 dB. Here, the CRB is the average of the CRBs obtained from 10^4 simulation runs. For the given N, L, and E_b/N_0, $\text{CRB}_{T1} \leq \text{CRB}_{T2} \leq \text{CRB}_{T3}$, where CRB_{Ti} denotes the CRB for Type i TSs. For $E_b/N_0 = 20$ dB, the MSEs were almost identical to the corresponding CRBs, as expected from (3) and (12), and the Type 1 TSs (proposed) performed somewhat better than the others. The CRBs and MSEs decreased monotonically as N increased.

Figs. 3 and 4 show the CRBs and MSEs of $\hat{\nu}$ against the SNR per bit when $(N, L) = (16, 8)$ and $(N, L) = (24, 8)$, respectively. Both i.i.d. Gaussian and GSM channels were considered. As in Fig. 2, $\text{CRB}_{T1} \leq \text{CRB}_{T2} \leq \text{CRB}_{T3}$. For high SNR values, say $E_b/N_0 \geq 20$ dB, the MSEs were almost identical to the corresponding CRBs. The ML estimate employing Type 1 TSs (proposed) performed better than the others; the estimate with Type 3 TSs (white) exhibited the worst performance; however, the difference between their MSE values was not significant. The advantage of Type 1 TSs over the others became more apparent when the SNR was low. In Fig. 3, where $(N, L) = (16, 8)$, the MSEs for Type 1 TSs were reasonably close to their corresponding CRBs, even for $E_b/N_0 = 5$ dB, while a considerable deviation was observed for the other types of TS. Although in Fig. 4 the MSEs for the Type 1 TSs also exhibited a deviation from the CRBs, the deviation tended to start at lower SNR

[2]In [10], a white TS is shown to be optimal in the minmax sense for an asymptotic CRB; applying the minmax method of (9) is not recommended due to its high computational complexity.

values, and the degree of deviation tended to be less than that for the other TSs. (In Fig. 4, the Type 3 TSs outperformed the Type 2 TSs for a low SNR.) It was observed that all the periodic (and complementarily periodic) TSs with $N = 2L$, shown in Table II, resulted in desirable MSE values for a low SNR that were close to the corresponding CRBs. Therefore, to improve the performance of the ML estimate for a low SNR, it is generally recommended to set $N = 2L$ so that periodic or complementarily periodic TSs can be found.

V. Conclusion

A necessary and sufficient condition for minimizing the UCRB for frequency estimation was derived, and a process for selecting TSs developed. The channel information required by the proposed approach is minimal: only the channel duration is needed. The proposed method was applied to selecting TSs for the ML frequency offset estimation, and useful binary TSs were found up to length 24. The advantage of the proposed TSs over some existing TSs was demonstrated through computer simulation.

The proposed method requires an exhaustive search, and its complexity exponentially increases with the TS length N. As such, this limits the use of the proposed method to selecting short TSs (say, $N < 30$ for binary TSs). Developing an algorithm for finding longer TSs and expanding the proposed TS selection to consider a joint ML frequency and channel estimation remain as future research topics.

Appendix A
Derivation of Lemma 1

Let $\mathbf{Q} = \text{diag}[\lambda_1, \ldots, \lambda_L]$, where λ_is are eigenvalues of $\mathbf{P}$ in (5). Then there exists a unitary matrix $\mathbf{U}$ such that $\mathbf{P} = \mathbf{U}^H\mathbf{Q}\mathbf{U}$ by the spectral theorem. Since $\mathbf{P}$ is positive definite, $\lambda_i > 0$ for each i. Notice that $\mathbf{Ux}$ is also a standard complex Gaussian random vector; therefore, $E[(\mathbf{x}^H\mathbf{P}\mathbf{x})^{-1}] = E[(\mathbf{x}^H\mathbf{U}^H\mathbf{Q}\mathbf{U}\mathbf{x})^{-1}] = E[(\mathbf{x}^H\mathbf{Q}\mathbf{x})^{-1}]$. It is straightforward to show that $\text{E}[(\mathbf{x}^H\mathbf{Q}\mathbf{x})^{-1}]$ is given by

$$E\left(\frac{1}{\mathbf{x}^H\mathbf{Q}\mathbf{x}}\right) = \int_{R^{2L}} \frac{\Phi(\mathbf{x})}{\lambda_1|x_1|^2 + \cdots + \lambda_L|x_L|^2}\, d\mathbf{x}. \quad \text{(A1)}$$

The convergence of (A1) is examined as follows. If λ_1 is the smallest positive number among λ_i, changing into spherical coordinates means that the RHS of (A1) is dominated by

$$\frac{(2\pi)^{-L}}{\lambda_1}\Sigma_{2L-1}\int_0^\infty e^{-r^2/2}r^{2L-3}\, dr \quad \text{(A2)}$$

where Σ_{2L-1} denotes the volume of the $(2L-1)$-dimensional sphere. The preceding integral converges if and only if $L > 1$.□

Appendix B
Derivation of Lemma 2

Let $\mathbf{H}f(\underline{\lambda})$ be the Hessian matrix of f at $\underline{\lambda} \in \Omega_m$; i.e., its (i,j)th entry is

$$\frac{\partial^2}{\partial\lambda_i\partial\lambda_j}f(\underline{\lambda}) = 2\int_{R^{2L}} \frac{|x_i|^2|x_j|^2\Phi(\mathbf{x})}{\left(\sum_{k=1}^L \lambda_k|x_k|^2\right)^3}\, d\mathbf{x}. \quad \text{(B1)}$$

Let $\mathbf{b} = [b_1, b_2, \ldots, b_L]^T \in R^L \setminus \{[0, 0, \ldots, 0]^T\}$. Then the inner product of $\mathbf{H}f(\underline{\lambda})\mathbf{b}$ with $\mathbf{b}$ is

$$\langle \mathbf{H}f(\underline{\lambda})\mathbf{b}, \mathbf{b}\rangle = 2\int_{R^{2L}} \frac{\left(\sum_{i=1}^L b_i|x_i|^2\right)^2\Phi(\mathbf{x})}{\left(\sum_{k=1}^L \lambda_k|x_k|^2\right)^3}\, d\mathbf{x} > 0. \quad \text{(B2)}$$

This shows that $\mathbf{H}f(\underline{\lambda})$ is positive definite for each $\underline{\lambda} \in (0,\infty)^L$. Now, [16, Theorem 3.6] implies that f is strictly convex in $(0,\infty)^L$. In particular, it is strictly convex in Ω_m. □

References

[1] A. Milewski, "Periodic sequences with optimal properties for channel estimation and fast start-up equalization," *IBM J. Res. Develop*, vol. 27, no. 5, pp. 426–431, Sep. 1983.

[2] S. N. Crozier, D. D. Falconer, and S. A. Mahmoud, "Least sum of squared errors (LSSE) channel estimation," *Proc. Inst. Elect. Eng. F*, vol. 138, no. 4, pp. 371–378, Aug. 1991.

[3] J. C. L. Ng, K. B. Letaief, and R. D. Murch, "Complex optimal sequences with constant magnitude for fast channel estimation initialization," *IEEE Trans. Commun.*, vol. 46, no. 3, pp. 305–308, Mar. 1998.

[4] C. Tellambura, Y. J. Guo, and S. K. Barton, "Channel estimation using aperiodic binary sequences," *IEEE Commun. Lett.*, vol. 2, no. 5, pp. 140–142, May 1998.

[5] C. Tellambura, M. G. Parker, Y. J. Guo, S. J. Shepherd, and S. K. Barton, "Optimal sequences for channel estimation using discrete Fourier transform techniques," *IEEE Trans. Commun.*, vol. 47, no. 2, pp. 230–238, Feb. 1999.

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

[7] S. Adireddy, L. Tong, and H. Viswanathan, "Optimal placement of training for frequency-selective block-fading channels," *IEEE Trans. Inf. Theory*, vol. 48, no. 8, pp. 2338–2353, Aug. 2002.

[8] X. Ma, G. B. Giannakis, and S. Ohno, "Optimal training for block transmission over doubly selective wireless fading channels," *IEEE Trans. Signal Process.*, vol. 51, no. 5, pp. 1351–1366, May 2003.

[9] H. Vikalo, B. Hassibi, B. Hochwald, and T. Kailath, "On the capacity of frequency-selective channels in training-based transmission schemes," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2572–2583, Sep. 2004.

[10] O. Besson and P. Stoica, "Training sequence selection for frequency offset estimation in frequency-selective channels," *Digital Signal Process.: Rev. J.*, vol. 13, no. 1, pp. 106–127, Jan. 2003.

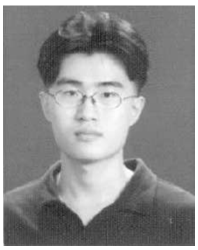
[11] P. Stoica and O. Besson, "Training sequence design for frequency offset and frequency-selective channel estimation," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1910–1917, Nov. 2003.

[12] M. Morelli and U. Mengali, "Carrier frequency estimation for transmissions over selective channels," *IEEE Trans. Commun.*, vol. 48, no. 9, pp. 1580–1589, Sep. 2000.

[13] M. Moeneclaey, "On the true and the modified Cramer-Rao bounds for the estimation of a scalar parameter in the presence of nuisance parameters," *IEEE Trans. Commun.*, vol. 46, no. 11, pp. 1536–1544, Nov. 1998.

[14] A. N. D'Andrea, U. Mengali, and R. Reggiannini, "The modified Cramer-Rao bound and its application to synchronization problems," *IEEE Trans. Commun.*, vol. 42, no. 2/3/4, pp. 1391–1399, Feb./Mar./Apr. 1994.

[15] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1990.
[16] W. Fleming, *Functions of Several Variables*, 2nd ed. New York: Springer-Verlag, 1977.



Young-Doo Kim (S'99) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Korea Advanced Institute of Science and Technology, Daejeon, Korea, in 1999, 2001, and 2006, respectively.

His primary research interests include synchronization, detection, and estimation for communication system and predistortion linearization of nonlinear power amplifiers for wireless applications.

Jae Kun Lim received the B.S. degree from Yonsei University, Seoul, Korea, in 1991 and the M.S. and Ph.D. degrees from Korea Advanced Institute of Science and Technology, Daejeon, in 1993 and 1998, respectively.

He was a Research Assistant Professor at the Center for High-Performance Integrated Systems, Daejeon, from 2001 to 2003, and then a full-time Lecturer at Mokpo National University, Mokpo, Korea, from 2003 to 2004. He is now an Assistant Professor with the Department of Computational Mathematics and Informatics, Hankyong National University, Anscong, Korea. His research interests include wavelet analysis, abstract frame theory, shift-invariant apaces, and information and communcation theory.

Changho Suh (M'04) received the B.S. and M.S. degrees in electrical engineering from Korea Advanced Institute of Science and Technology, Daejeon, Korea, in 2000 and 2002, respectively.

Since 2002, he has been with the Department of Telecommunication R&D Center, Samsung Electronics Co., Ltd., Kyonggido, Korea, where he is working on a fourth-generation communications project as a Research Engineer. His research interests span the areas of wireless communications and signal processing, resource management in multicarrier wireless network, and sequence and coding theory. Currently, he is focusing on the resource management in wireless *ad hoc* and sensory networks.

Yong H. Lee (S'81–M'84–SM'98) was born in Seoul, Korea, on July 12, 1955. He received the B.S. and M.S. degrees in electrical engineering from Seoul National University in 1978 and 1980, respectively, and the Ph.D. degree in electrical engineering from the University of Pennsylvania, Philadelphia, in 1984.

From 1984 to 1988, he was an Assistant Professor with the Department of Electrical and Computer Engineering, State University of New York, Buffalo. Since 1989, he has been with the Department of Electrical Engineering and Computer Science, Korea Advanced Institute of Science and Technology, Daejeon, Korea, where he currently is a Professor and the Dean of Engineering College. His research activities are in the area of commmunication signal processing, which includes synchronization, estimation, detection, interference cancellation, and efficient filter design for CDMA, TDMA, and orthogonal frequency-division multiplexing systems. He is also interested in designing and implementing transceivers.