

# Transactions Papers

## Pilot Designs for Consistent Frequency-Offset Estimation in OFDM Systems

Yinghui Li, *Student Member, IEEE*, Hlaing Minn, *Member, IEEE*, Naofal Al-Dhahir, *Senior Member, IEEE*, and A. Robert Calderbank, *Fellow, IEEE*

**Abstract**—This paper presents pilot designs for consistent frequency-offset estimation of orthogonal frequency-division multiplexing systems in frequency-selective fading channels. We describe two design approaches, namely, consistency in the probabilistic sense and absolute consistency. Existing preambles and pilot designs in the literature do not guarantee the absolute consistency. We derive general criteria for both approaches, present sufficient conditions on the pilot structures over the maximum carrier frequency offset (CFO) estimation range (half of the sampling rate), and derive simple pilot designs satisfying these conditions. We also extend the sufficient conditions to any arbitrary but fixed CFO estimation range, and present some generalized design patterns. Furthermore, the CFO estimation performances of distinct consistent pilot designs can be quite different at moderate or low signal-to-noise ratio (SNR) due to different statistics of outliers which also yields a link failure. We develop efficient pilot-design criteria that provide both consistency and robustness against outliers at moderate-to-low SNR. Our consistent pilot designs facilitate flexible and economical implementation, while our robust pilot designs enable wireless links with less outage and better resilience.

**Index Terms**—Bose–Chaudhuri–Hocquengem (BCH) code, consistency, cyclic-difference set, frequency offset, m-sequence, maximum-likelihood (ML) estimation, orthogonal frequency-division multiplexing (OFDM), outlier, pilot design, robustness.

### I. INTRODUCTION

CARRIER frequency offset (CFO) estimation plays a crucial role in establishing a reliable wireless communication link. Preamble or pilot-aided CFO estimation is adopted in all

Paper approved by C. Tepedelenlioglu, the Editor for Transmission Systems of the IEEE Communications Society. Manuscript received September 4, 2005; revised June 8, 2006 and October 13, 2006. The work of H. Minn was supported in part by the E. Jonsson School Research Excellence Initiative, the University of Texas at Dallas. The work of N. Al-Dhahir was supported in part by the Texas Advanced Technology Program (ATP) under Contract 009741-0023-2003, in part by the National Science Foundation under Contracts CCF 04-30654 and DMS 05-28010, and in part by a gift from Texas Instruments Inc. The work of A. R. Calderbank was supported in part by National Science Foundation under Contract DMS 05-28010. This paper was presented in part at the IEEE International Conference on Communications, Istanbul, Turkey, June 2006, and in part at IEEE MILCOM, Washington, DC, October 2006.

Y. Li, H. Minn, and N. Al-Dhahir are with the Department of Electrical Engineering, University of Texas at Dallas, Richardson, TX 75083-0688 USA (e-mail: yinghui.li@utdallas.edu; hlaing.minn@utdallas.edu; aldhahir@utdallas.edu).

A. R. Calderbank is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: calderbk@princeton.edu).

Digital Object Identifier 10.1109/TCOMM.2007.896105

current wireless communication systems (e.g., 2G, 3G, wireless PAN/LAN/MAN, etc.) to achieve a reliable estimation performance with low latency and complexity. The maximum absolute value of the CFO which can be handled by an estimator is one-half of the sampling rate. In most wireless systems and existing works in the literature (e.g., [1]–[9]), handling such a large CFO is not necessary, and hence, the preambles are designed for a smaller estimation range. An important issue that most of the previous works overlooked is the CFO estimation consistency, defined as the unambiguousness of the estimation metric trajectory within the considered estimation range in the absence of noise, regardless of the channel impulse response (CIR). For consumer-related applications, inconsistency of the estimation under certain channel conditions may be tolerable, but consistency will still be a desirable feature if the associated cost is affordable. For emergency and other critical wireless systems (e.g., flight control, space shuttle control, military combat scenarios, etc.), the consistency should not be compromised. As an example, consider a scenario where a person is in an emergency and trying to make a 911 call. Suppose the corresponding channel response is not changing and happens to yield an inconsistent estimate. Then, a communication link will not be established due to a complete loss in frequency synchronization. This example highlights the importance of estimation consistency for emergency scenarios.

Only a few works in the literature addressed the consistency issue in CFO estimation. [10] was the first to address the consistency issue for blind CFO estimation in orthogonal frequency-division multiplexing (OFDM) systems. Two approaches, using null tones with distinctive spacings and randomly hopping the null tones from block to block, were proposed for consistent estimation. Consistency issues for null-tones-based CFO estimators were also addressed in [11] and [12]. On the other hand, [13] (following [10]) was the first in addressing the consistency for *pilot-based* CFO estimation in OFDM systems. Distinctively spaced pilot tones were presented for its proposed CFO estimator. We note that the maximum-likelihood estimator (MLE) in [14] gives slightly better estimation performance than the estimator in [13] (see Fig. 1). However, pilot design for consistency in the MLE has not been addressed, and is one of the main topics of this paper. Furthermore, the distinctive pilot tones used in [13] may not guarantee the consistency, as will be shown later.

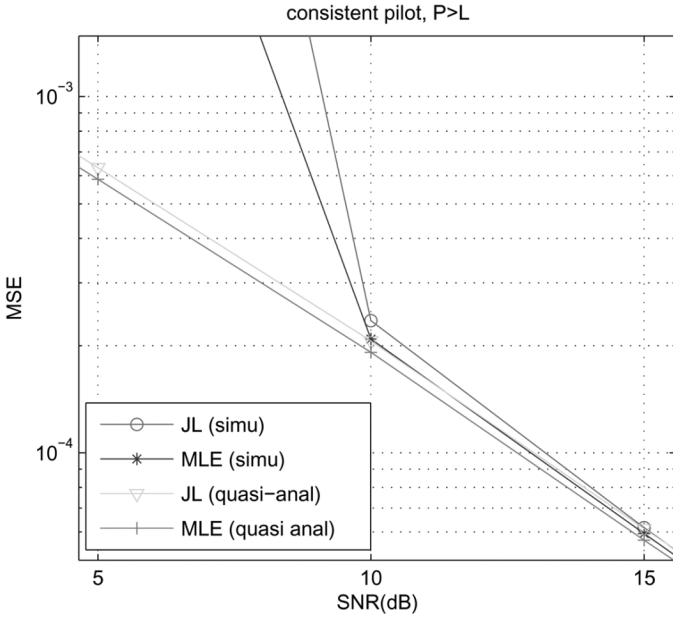


Fig. 1. Mean-square error comparison of the MLE and the JL estimator.

Hence, this paper will also address pilot design ensuring consistency of the estimator in [13].

For critical communications systems (e.g., those for emergency or military), consistency over the entire CFO range of half-sampling rate may be desirable. But for consumer-related wireless systems, consistency over a smaller CFO range may be preferred, due to complexity and cost concerns. Due to their own merits, we address both pilot designs with consistency for the entire CFO range and the designs for a smaller range.

Another important problem that we will answer in this paper is, “Which consistent pilot design will give better CFO estimation performance?” We observe that estimation performances of consistent pilots can be significantly different at moderate-to-low signal-to-noise ratio (SNR), mainly due to the distinct statistics of outliers which cause large CFO estimation errors resulting in link failures. An explanation is that some consistent pilots may have an estimation metric trajectory with sidelobe peaks comparable to the mainlobe peak for some channel responses yielding less robust performance at low SNR. Hence, both robust and consistent pilot designs are crucial for emergency-related or other critical communication systems.

The scope of this paper encompasses pilot designs for consistent and robust CFO estimation of OFDM systems in frequency-selective multipath fading channels. Our main contributions are summarized as follows. First, we derive general necessary and sufficient criteria for the consistency of the MLE and the estimator in [13]. We propose two consistency criteria over the maximum CFO estimation range (half of the sampling rate): one for the absolute consistency (or simply denoted as consistency), and the other for consistency in the probabilistic sense. Second, in relating the estimators’ consistency criteria to the pilot design, we present sufficient conditions on the pilot structures to yield consistency over the maximum CFO estimation range, and derive simple pilot designs satisfying these conditions. Third, we extend the sufficient conditions to

an arbitrary smaller CFO estimation range, and present several generalized design patterns. Finally, we develop efficient design criteria for robustness of the consistent pilots against outliers. Note that the optimal and consistent pilot design for both CFO and channel estimation is an ambitious open problem, where the optimality requirements of one task may be conflicting with those of the other. It is out of the scope of this paper. Some initial attempts along these lines are discussed in [15] and [16].

The rest of this paper is organized as follows. Section II describes the signal model and the CFO estimation algorithms. Sections III–VII present our contributions mentioned above. Section VIII provides simulation results and discussions, and Section IX concludes the paper.

*Notations:* A bold small (capital) letter represents a column vector (matrix). The superscripts  $*$ ,  $T$ ,  $H$  represent the conjugate, the transpose, and the conjugate transpose operations, respectively. The all-one (all-zero) column vector of length  $k$ , the  $k \times m$  all-zero matrix, and the  $k \times k$  identity matrix are denoted by  $\mathbf{1}_k$  ( $\mathbf{0}_k$ ),  $\mathbf{0}_{k \times m}$ , and  $\mathbf{I}_k$ , respectively. The  $i$ th column of  $\mathbf{I}_N$  is denoted by  $\mathbf{e}_i$ , and  $\text{diag}\{\mathbf{x}\}$  represents a diagonal matrix whose diagonal elements are defined by  $\mathbf{x}$ . The  $l$ -cyclic-shifted version of  $\mathbf{c}$  is denoted by  $\mathbf{c}^{(l)}$ .  $\lfloor X \rfloor$  denotes the largest integer less than or equal to  $X$ , while  $\lceil X \rceil$  represents the smallest integer greater than or equal to  $X$ .

## II. SIGNAL MODEL AND FREQUENCY OFFSET ESTIMATION

We consider a discrete-time low-pass-equivalent complex baseband OFDM system with  $N$  subcarriers. The time-domain preamble/training sample is given by

$$s_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} c_n e^{j2\pi nk/N} \quad (1)$$

$$k = -N_g, -N_g + 1, \dots, N - 1$$

where  $N_g$  is the number of cyclic prefix (CP) samples (generally larger than the number of sample-spaced channel taps  $L$ ) and  $\{c_n\}$  are pilot tones (some of them can be zeros). Denote the indices of nonzero pilot tones and null tones by  $\{t_k : k = 1, \dots, P\}$  and  $\{n_k : k = 1, \dots, N - P\}$ , respectively (i.e.,  $c(t_k) \neq 0$  and  $c(n_k) = 0$ ), where  $\{t_k\} \cap \{n_k\} = \emptyset$  and  $\{t_k\} \cup \{n_k\} = \{0, 1, \dots, N - 1\}$ . In the presence of a normalized (by the subcarrier spacing) CFO  $v$ , the received training signal vector is given by

$$\mathbf{r} = \mathbf{\Gamma}(v)\mathbf{S}\mathbf{h} + \mathbf{w} \quad (2)$$

where

$$\mathbf{h} = [h_0, h_1, \dots, h_{L-1}]^T \quad (3)$$

$$\mathbf{w} = [w_0, w_1, \dots, w_{N-1}]^T \quad (4)$$

$$\mathbf{\Gamma}(v) = \text{diag}\{1, e^{j2\pi v/N}, \dots, e^{j2\pi(N-1)v/N}\} \quad (5)$$

$$[\mathbf{S}]_{k,n} = s_{k-n}, \quad 0 \leq k \leq N - 1, \quad 0 \leq n \leq L - 1. \quad (6)$$

In the above equations,  $\{w_n\}$  are independent and identically distributed zero-mean circularly symmetric complex Gaussian noise samples, each having a variance of  $\sigma_n^2$ , and  $\{h_k\}$  are the equivalent sample-spaced channel tap gains (including the

transmit and receive filters) assumed to remain constant during the training block. Denote the unitary  $N$ -point fast Fourier transform (FFT) matrix by  $\mathbf{F} = [\mathbf{f}_0, \mathbf{f}_1, \dots, \mathbf{f}_{N-1}]$ , where  $\mathbf{f}_k = [1, e^{-j2\pi k/N}, \dots, e^{-j2\pi k(N-1)/N}]^T / \sqrt{N}$ . Define

$$\mathbf{\Theta} = [\mathbf{e}_{t_1}, \mathbf{e}_{t_2}, \dots, \mathbf{e}_{t_P}, \mathbf{e}_{n_1}, \dots, \mathbf{e}_{n_{N-P}}] \quad (7)$$

$$\mathbf{F}_L = [\mathbf{f}_0, \mathbf{f}_1, \dots, \mathbf{f}_{L-1}] \quad (8)$$

$$\mathbf{F}_P = [\mathbf{f}_{t_1}, \mathbf{f}_{t_2}, \dots, \mathbf{f}_{t_P}] = \mathbf{F}\mathbf{\Theta}[\mathbf{I}_P, \mathbf{0}_{P \times (N-P)}]^T \quad (9)$$

$$\mathbf{F}_Z = [\mathbf{f}_{n_1}, \mathbf{f}_{n_2}, \dots, \mathbf{f}_{n_{(N-P)}}] = \mathbf{F}\mathbf{\Theta}[\mathbf{0}_{(N-P) \times P}, \mathbf{I}_{N-P}]^T \quad (10)$$

$$\mathbf{c} = [c_0, c_1, \dots, c_{N-1}]^T \quad (11)$$

$$\mathbf{C} = \text{diag}\{\mathbf{c}\} \quad (12)$$

$$\tilde{\mathbf{h}} = [H_0, H_1, \dots, H_{N-1}]^T = \mathbf{F}_L \mathbf{h} \quad (13)$$

$$\mathbf{H} = \text{diag}\{\tilde{\mathbf{h}}\}. \quad (14)$$

Then, the time-domain training signal matrix  $\mathbf{S}$  and received vector  $\mathbf{r}$  can be expressed as

$$\mathbf{S} = \sqrt{N} \mathbf{F}^H \mathbf{C} \mathbf{F}_L \quad (15)$$

$$\mathbf{r} = \sqrt{N} \mathbf{\Gamma}(v) \mathbf{F}^H \mathbf{C} \mathbf{F}_L \mathbf{h} + \mathbf{w} = \sqrt{N} \mathbf{\Gamma}(v) \mathbf{F}^H \mathbf{H} \mathbf{c} + \mathbf{w}. \quad (16)$$

The MLE<sup>1</sup> from [14] for the normalized CFO  $v$  is given by

$$\hat{v} = \arg \max_{\tilde{v}} \{g_1(\tilde{v})\} \quad (17)$$

where

$$g_1(\tilde{v}) = \mathbf{r}^H \mathbf{\Gamma}(\tilde{v}) \mathbf{B}_1 \mathbf{\Gamma}^H(\tilde{v}) \mathbf{r}. \quad (18)$$

In the above equation,  $\mathbf{B}_1$  is an  $N \times N$  projection matrix given by

$$\mathbf{B}_1 = \mathbf{S}(\mathbf{S}^H \mathbf{S})^{-1} \mathbf{S}^H. \quad (19)$$

On the other hand, the CFO estimator from [13] is given by

$$\hat{v} = \arg \max_{\tilde{v}} \{g_2(\tilde{v})\} \quad (20)$$

where

$$g_2(\tilde{v}) = \mathbf{r}^H \mathbf{\Gamma}(\tilde{v}) \mathbf{B}_2 \mathbf{\Gamma}^H(\tilde{v}) \mathbf{r} \quad (21)$$

and

$$\mathbf{B}_2 = \mathbf{F}_P^* \mathbf{F}_P^T. \quad (22)$$

The mean-square error (MSE) of the MLE in (17) is smaller than that of the estimator in (20) when the number of nonzero pilot tones  $P$  is larger than the number of channel taps  $L$ . A proof is given in Appendix A.

Note that both estimators were developed based on maximizing the likelihood function for the joint estimation of the frequency offset and the channel. Both estimators replace the channel vector in the likelihood function with the maximum-likelihood (ML) channel estimate. The difference is that the MLE in [14] used a time-domain model and exploiting the

channel's limited delay spread ( $L$  taps) was embedded in its time-domain ML channel estimate, while the estimator in [13] used a frequency-domain model and the channel's limited delay spread property was not exploited in its frequency-domain channel estimation. Hence, the MLE in [14] requires the knowledge of  $L$ , the maximum number of sample-spaced equivalent channel taps, which can be set to an appropriate upper bound in practice. The limited delay spread translates into frequency correlation property of the channel. Hence, when all  $L$  channel taps are identifiable in the channel estimation (which requires that  $P \geq L$ ), the MLE in [14] outperforms the estimator in [13] since it uses the frequency correlation property of the channel. On the other hand, the latter has a complexity advantage over the former.

### III. THE ESTIMATOR CONSISTENCY CONDITION

The estimator consistency condition can be stated as follows. In the absence of noise, there is only one  $\tilde{v}$  that maximizes the estimation metric  $g_i(\tilde{v})$ , and it is at  $\tilde{v} = v$  for any  $\mathbf{h} \neq \mathbf{0}_L$ . Define

$$G_i(\Delta) = g_i(v) - g_i(\tilde{v}) \quad (23)$$

where  $\Delta = v - \tilde{v}$  and the range of  $\Delta$  is  $(v - N/2, v + N/2)$ . From the MLE principle,  $g_i(\tilde{v})$  is maximized at  $\tilde{v} = v$ , so  $G_i(\Delta)$  is nonnegative. Then, the estimator consistency condition is given as

$$G_i(\Delta) = 0 \text{ if and only if } \Delta = 0 \quad \forall \mathbf{h} \neq \mathbf{0}_L. \quad (24)$$

For the MLE, we have

$$G_1(\Delta) = \mathbf{h}^H \mathbf{S}^H \mathbf{B}_1 \mathbf{S} \mathbf{h} - \mathbf{h}^H \mathbf{S}^H \mathbf{\Gamma}^H(\Delta) \mathbf{B}_1 \mathbf{\Gamma}(\Delta) \mathbf{S} \mathbf{h}. \quad (25)$$

By singular value decomposition, we have  $\mathbf{S} = \mathbf{U} \mathbf{\Sigma}_S \mathbf{V}^H$  where  $\mathbf{U} = [\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{N-1}]$  is the unitary matrix containing the eigenvectors of  $\mathbf{S} \mathbf{S}^H$ ,  $\mathbf{V}$  is the unitary matrix containing the eigenvectors of  $\mathbf{S}^H \mathbf{S}$ ,  $\mathbf{\Sigma}_S$  is the  $N \times L$  diagonal matrix with nonincreasing singular values of  $\mathbf{S}$ , and  $\mu = \text{rank}(\mathbf{S})$ . Then, we obtain

$$\mathbf{B}_1 = \mathbf{U} \mathbf{\Sigma}_S (\mathbf{\Sigma}_S^H \mathbf{\Sigma}_S)^{-1} \mathbf{\Sigma}_S^H \mathbf{U}^H = \mathbf{U} \mathbf{\Sigma}_B \mathbf{U}^H \quad (26)$$

$$\mathbf{\Sigma}_B = \text{diag}\{\mathbf{1}_\mu^T, \mathbf{0}_{N-\mu}^T\} \quad (27)$$

where  $(\cdot)^{-1}$  is the matrix inverse (treated as a pseudoinverse for a rank-deficient matrix).<sup>2</sup> After substituting (26) into (25), we obtain

$$G_1(\Delta) = \mathbf{z}_1^H(\Delta) \mathbf{z}_1(\Delta) \quad (28)$$

where

$$\mathbf{z}_1(\Delta) = (\mathbf{I} - \mathbf{\Sigma}_B) \mathbf{U}^H \mathbf{\Gamma}(\Delta) \mathbf{U} \mathbf{\Sigma}_S \mathbf{V}^H \mathbf{h}. \quad (29)$$

From (28), we can easily see that  $G_1(\Delta) = 0$  if and only if  $\mathbf{z}_1(\Delta) = \mathbf{0}_N$ . Then, the consistency condition requires that

<sup>2</sup>For the rank-deficient case (i.e.,  $P < L$ ), the estimator in (18) is no longer the MLE, but for convenience, we use the same acronym (MLE). For  $P < L$ , it has the same estimation performance as the estimator in [13], as proved in Appendix A.

<sup>1</sup>It is based on a joint ML estimation of the CIR and the CFO.

$\mathbf{z}_1(\Delta) = \mathbf{0}_N$  only at  $\Delta = 0$  for any  $|\tilde{v}| < N/2$ . In other words,  $\mathbf{z}_1(\Delta) \neq \mathbf{0}_N$  for any  $|\tilde{v}| < N/2$  except  $\tilde{v} = v$ . This condition can be achieved for any  $\mathbf{h}$  ( $\neq \mathbf{0}_L$ ) if the following necessary and sufficient condition is satisfied:

$$\text{rank}((\mathbf{I} - \Sigma_{\mathbf{B}})\mathbf{U}^H\mathbf{\Gamma}(\Delta)\mathbf{U}\Sigma_{\mathbf{S}}\mathbf{V}^H) = L. \quad (30)$$

By using (27), the condition in (30) becomes

$$P \geq L \ \& \ \text{rank}(\mathbf{U}_2^H\mathbf{\Gamma}(\Delta)\mathbf{U}_1) = L \quad \forall \Delta \neq 0 \quad (31)$$

where

$$\mathbf{U}_1 = [\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{L-1}] \quad (32)$$

$$\mathbf{U}_2 = [\mathbf{u}_L, \mathbf{u}_{L+1}, \dots, \mathbf{u}_{N-1}]. \quad (33)$$

For the estimator of [13], we have

$$G_2(\Delta) = \mathbf{z}_2^H(\Delta)\mathbf{z}_2(\Delta) \quad (34)$$

where

$$\mathbf{z}_2(\Delta) = \sqrt{N}\mathbf{F}_Z^T\mathbf{\Gamma}(\Delta)\mathbf{F}^H\mathbf{H}\mathbf{c} = \sqrt{N}\mathbf{F}_Z^T\mathbf{\Gamma}(\Delta)\mathbf{F}^H\mathbf{C}\mathbf{F}_L\mathbf{h}. \quad (35)$$

Then, the consistency condition requires that  $\mathbf{z}_2(\Delta) \neq \mathbf{0}_N$  for any  $|\tilde{v}| < N/2$  except  $\tilde{v} = v$ . This condition can be achieved for any  $\mathbf{h}$  ( $\neq \mathbf{0}_L$ ) if we have the following necessary and sufficient condition:

$$\text{rank}(\mathbf{F}_Z^T\mathbf{\Gamma}(\Delta)\mathbf{F}^H\mathbf{C}\mathbf{F}_L) = L \quad \forall \Delta \neq 0 \quad (36)$$

which can be satisfied by the following sufficient condition:

$$(N - P) \geq P \ \& \ P \geq L \ \& \ \text{rank}(\mathbf{F}_Z^T\mathbf{\Gamma}(\Delta)\mathbf{F}_P^*) \geq L \quad \forall \Delta \neq 0. \quad (37)$$

#### IV. PROPOSED PILOT DESIGNS

We start in this section by presenting pilot design conditions sufficient to satisfy the consistency conditions in (31) for the MLE and in (37) for the estimator of [13]. Then, simple consistent pilot designs are developed for both estimators. Denote the indices of nonzero pilot tones and null tones of  $\mathbf{c}^{(l)}$  by  $\{t_k^{(l)} : k = 1, \dots, P\}$  and  $\{n_k^{(l)} : k = 1, \dots, N - P\}$ , respectively, where  $t_k^{(l)} = (t_k + l) \bmod N$ ,  $n_k^{(l)} = (n_k + l) \bmod N$ ,  $\{t_k^{(l)}\} \cap \{n_k^{(l)}\} = \emptyset$ , and  $\{t_k^{(l)}\} \cup \{n_k^{(l)}\} = \{0, 1, \dots, N - 1\}$ .

Then, a sufficient pilot design condition for the consistency of the MLE is that  $(N - P) \geq P$ , and for any cyclic-shifting distance  $l \in \{1, 2, \dots, N - 1\}$ ,  $\mathbf{c}^{(l)}$  always has at least  $L$  nonzero pilot tones located at the null-tone indices of the original unshifted pilot vector  $\mathbf{c}$ .

In other words,  $(N - P) \geq P$  and the cardinality of  $\{t_k^{(l)}\} \cap \{n_k\}$ , for any  $l \in \{1, 2, \dots, N - 1\}$ , is always greater than or equal to  $L$ . The above condition implicitly requires that  $P > L$ . The proof is provided in Appendix B.

Define  $\mathbf{p}$  as the pilot-tone location vector whose elements are given by

$$[\mathbf{p}]_k = \begin{cases} 1, & k \in \{t_k\} \\ 0, & \text{otherwise.} \end{cases} \quad (38)$$

Then the above pilot design sufficient condition for the MLE can be expressed as

$$(\mathbf{1}_N - \mathbf{p})^T \mathbf{p}^{(l)} \geq L \quad \forall l \in \{1, \dots, N - 1\} \quad (39)$$

$$\Rightarrow \phi(l) \leq P - L \quad \forall l \in \{1, \dots, N - 1\} \quad (40)$$

where  $\phi(l) = \mathbf{p}^T \mathbf{p}^{(l)}$  is the *periodic autocorrelation* function of  $\mathbf{p}$ . In the following, we develop simple pilot designs satisfying the above sufficient condition. We link the condition in (39), or equivalently (40), to the properties of cyclic codes. We first describe a general design based on a cyclic code which is followed by specific designs based on a binary Bose–Chaudhuri–Hocquengem (BCH) code, an m-sequence, a difference-set cyclic code, a cyclic difference-set, and a distinctive pilot spacing.

##### A. Design Based on a Cyclic Code

Except the all-zero codeword, all other codewords of a cyclic code are just cyclic-shifted versions of each other. This property is consistent with our considered scenario of  $\mathbf{p}^{(l)}$ . The Hamming distance of two binary codewords  $\mathbf{p}$  and  $\mathbf{p}^{(l)}$  with the Hamming weight  $P$  is given by

$$d_H(\mathbf{p}, \mathbf{p}^{(l)}) = 2(P - \mathbf{p}^T \mathbf{p}^{(l)}). \quad (41)$$

From (40) and (41), the required consistency condition becomes

$$d_H(\mathbf{p}, \mathbf{p}^{(l)}) \geq 2L. \quad (42)$$

The Hamming distance between any two codewords in a binary code is bounded below by the minimum distance of the code. Then, a simple pilot design satisfying the consistency condition for the MLE is given by any codeword (except the all-zero codeword) of a binary cyclic code with the minimum Hamming distance  $d_{\min} \geq 2L$  and the codeword length  $N_1 \leq N$ . If  $N_1 < N$ , the codeword length adjustment by means of inserting zeros to give a length- $N$  vector is discussed in Section IV-G. The binary cyclic codes also contain binary BCH codes, maximum-length codes, and difference-set cyclic codes. In the following, we will describe these more specific pilot designs.

##### B. Design Based on a Binary BCH Code

A  $(n, k, t)$  binary BCH code satisfies the following relationships:  $n = 2^m - 1$ ;  $n - k \leq mt$ ;  $d_{\min} = 2t + 1$ , where  $m (\geq 3)$  and  $t$  are any positive integers. Any codeword (except the all-zero codeword) of a binary BCH code with  $t \geq (2L - 1)/2$  and  $n \leq N$  (inserted with zeros if  $n < N$ , see Section IV-G) can be used as a consistent pilot location vector. The coefficients of generator polynomials for several binary BCH codes (with lengths up to 1023) can be easily found in the coding theory literature (e.g., [17]).

### C. Design Based on a Maximum-Length Code or an m-sequence

A maximum-length code is generated by an  $m$ -stage shift register and satisfies the following relationship:  $(n, k) = (2^m - 1, m)$  and  $d_{\min} = 2^{m-1}$ . Nontrivial maximum-length codes exist for any positive integer  $m (\geq 3)$ . In fact, they are maximum-length sequences (m-sequences), and shift register connections for generating m-sequences can be easily found in the literature (e.g., [18]). An m-sequence  $\mathbf{p}$  of length  $N_1 = 2^m - 1$  has the Hamming weight of  $P = (N_1 + 1)/2$ . The periodic autocorrelation function of a bipolar ( $\pm 1$  or binary phase-shift keying (BPSK)-modulated) m-sequence, denoted by  $\Phi(l)$ , can be related to  $\phi(l)$  as

$$\begin{aligned} \Phi(l) &= (2\mathbf{p} - \mathbf{1}_{N_1})^T (2\mathbf{p}^{(l)} - \mathbf{1}_{N_1}) = 4\phi(l) - 4P + N_1 \\ &= 4\phi(l) - 2P - 1. \end{aligned} \quad (43)$$

It is well known that  $\Phi(l)$  takes on only two values, i.e.,  $\Phi(l) = N_1$  for  $l = iN_1$  ( $i = \text{integer}$ ) and  $\Phi(l) = -1$  otherwise. The second value translates into  $\phi(l \neq iN_1) = P/2$ . Incorporating the condition from (40) together with  $P = (N_1 + 1)/2$  gives a simple consistent pilot design based on an m-sequence with length  $N_1 \geq 4L - 1$  which is inserted with zeros, if necessary, to form a length- $N$  pilot location vector as described in Section IV-G.

### D. Design Based on a Difference-Set Cyclic Code

A perfect difference set of order  $q$  and modulus  $n = q(q+1)+1$  is defined as  $Q = [l_0, l_1, \dots, l_q]$ ,  $0 \leq l_0 < l_1 < \dots < l_q \leq q(q+1)$ , where no two of the  $q(q+1)$  ordered differences  $\{l_j - l_i : j \neq i\}$  modulo  $n$  are identical. Let  $Q(x) = x^{l_0} + x^{l_1} + \dots + x^{l_q}$  and  $h(x) = \text{GCD}\{Q(x), x^n + 1\}$ . Then, the difference-set cyclic code is generated by the generator polynomial  $g(x) = (x^n + 1)/h(x)$ . For  $q = 2^s$ , where  $s$  is any positive integer,  $n = 2^{2s} + 2^s + 1$ ,  $n - k = 3^s + 1$ , and the  $(n, k)$  difference-set cyclic code has the minimum Hamming distance of  $d_{\min} = 2^s + 2$ . A list of binary difference-set cyclic codes can be found in [17] and [19]. The consistent pilot design can be based on a codeword of a difference-set code with length  $n (\leq N)$  and  $d_{\min} \geq 2L$  (inserted with zeros, if  $n < N$ , to obtain a length- $N$  vector, as discussed in Section IV-G).

### E. Design Based on a Cyclic Difference Set

Alternatively, the consistent pilot-tone location vector can be directly based on a properly chosen perfect difference set<sup>3</sup> with  $q = P - 1$  and  $n \leq N$  (i.e., a cyclic difference set with  $n \leq N$ ,  $k = P$ , and  $t = 1$ ). An  $(n, k, t)$  cyclic difference set is a subset  $\{l_0, l_1, \dots, l_{k-1}\}$  of the integers modulo  $n$ , such that each element of  $\{1, 2, \dots, n-1\}$  can be represented as  $\{l_j - l_i : j \neq i\}$  modulo  $n$  in exactly  $t$  ways. For  $t = 1$ , the distinctive difference property of a properly chosen cyclic difference set (which yields the distinctive pilot spacings) satisfies our consistent pilot design criterion. Cyclic difference sets exist for certain lengths  $n$ . If  $N = n$ , the  $(t = 1)$  cyclic difference set with  $k = P > L$  can be used as the pilot location vector. If a cyclic difference set

<sup>3</sup>The condition  $n = q(q+1)+1$  in the definition of a perfect difference set is not necessary for our pilot design.

with  $n = N$  does not exist, we can use a  $(t = 1)$  cyclic difference set with  $n \leq N/2$  (which guarantees that all the differences  $\{l_j - l_i : j \neq i\}$  modulo  $N$  are distinctive) and  $k = P > L$ . A list of  $(n, k, t)$  cyclic difference sets can be found in [20].

### F. Design Based on Distinctive Pilot Spacings

More general (cyclic) distinctive pilot spacings, where the condition  $n = q(q+1)+1$  is not necessary, can alternatively be easily generated by computer.<sup>4</sup> Note that [13] used a distinctive pilot spacing but the condition  $P > L$  was not imposed, and hence, the consistency condition in (24) is not guaranteed. Furthermore, the pilot location designs from [13] (i.e.,  $\{t_k\} = \{2^{k-1}\}$  and  $\{t_k\} = \{2^{k+2}-7\}$ ) may not satisfy  $P > L$  for practical system parameters  $N$  and  $L$ . For example, an IEEE 802.11a system ( $N = 64$ ) can easily have  $L > 5$ , for which the pilot location designs used in [13] will not ensure consistency.

### G. Length Adjustment in Consistent Pilot Designs

To obtain length- $N$  consistent pilot vectors, the above designs first construct length- $N_1 (\leq N)$  consistent pilot vectors. For typical OFDM system,  $N$  is a power of 2 (we also assume so in this paper), while  $N_1$  can vary depending on the design. Note that appending zeros to get a length- $N$  vector may not always guarantee the consistency. In this section, we present several zero-padding designs for converting length- $N_1$  consistent pilot vectors into length- $N$  consistent pilot vectors.

- 1) If the number of padded zeros are equal to  $kP$  where  $k = 1, 2, \dots$ , then  $k$  zeros can be inserted between every two adjacent nonzero pilot tones.
- 2) If  $N_1 \leq N/2$ ,  $N - N_1$  zeros can be inserted between any two consecutive tones.
- 3) If  $N/2 < N_1 < N$  and  $N - N_1 \neq kP$ , we can either find the zero-padding structures by analyzing the specific design structures, or find where to insert  $N - N_1$  zeros by numerically evaluating (39). For example, we present several zero-padding structures for some specific design patterns as follows.

3a) Distinctively spaced pilot structure: In order to preserve the consistency,  $N - N_1$  zeros can be inserted between the two adjacent nonzero pilots with the largest cyclic distance to each other among all the adjacent nonzero pilot pairs.

3b) m-sequence-based pilot design: First, find the value of  $N_L$ , which is the minimum length of the cyclically sliding window, which always contains at least  $L$  nonzero pilot tones, by cyclically searching the original m-sequence vector. If  $N - N_1 \geq N_L$ , insert zeros between any two adjacent pilot tones: 1) until the new pilot vector length reaches  $N$ ; or 2) until the new sequence length is  $N_2$ , which satisfies  $N - N_2 = kP$ ,  $k = 1, 2, \dots$ , and then insert  $k$  zeros between every two adjacent nonzero pilot tones and finally obtain the length- $N$  pilot vector. If  $N - N_1 < N_L$ , we can find where to insert  $N - N_1$  zeros by numerically evaluating (39).

<sup>4</sup>By checking the distinctive spacing property for the pilot spacings  $1, 2, 3, \dots$ , iteratively.

3c) BCH-code-based pilot design: Since  $N_1 = 2^m - 1$  for BCH codes and  $N = 2^k$ , if  $k > m$ , we can insert  $N - N_1$  zeros between any two adjacent pilot tones. If  $k = m$ , we can find where to insert one zero by numerically evaluating (39).

#### H. Pilot Designs for the Estimator in [13]

Fortunately, our proposed pilot designs described above can be directly applied to the estimator of [13], since the sufficient pilot design condition for the MLE is also sufficient to satisfy the consistency condition in (37) for the estimator of [13]. The proof is provided in Appendix B.

#### I. Spectrum Shaping

The pilot tone location vector obtained from any of the above designs is then appropriately cyclic-shifted to satisfy the spectrum and system requirement (i.e., no pilot tones at DC and band edges). For example, suppose the obtained nonzero pilot tone indices are 0, 1, 3, 7, 12, 20, 30 for  $N = 64$  and  $L < 7$ . Then, they can be cyclic-shifted to 5, 6, 8, 12, 17, 25, 35 which give null tones at DC, 5 subcarriers at the lower band edge, and 4 subcarriers at the higher band edge. Also note that the proposed simple pilot designs are just a subset of the consistent pilot signals satisfying (31) for the MLE and (36) for the estimator of [13]. If necessary, (31) or (36) can be directly used in the design.

#### V. CONSISTENCY IN THE PROBABILISTIC SENSE

The consistency discussed in the previous section holds for any  $\mathbf{h}$  ( $\neq \mathbf{0}_L$ ). Hence, the conditions and the pilot designs from the previous section guarantee absolute consistency. In this section, we consider consistency in the probabilistic sense (i.e., for some  $\mathbf{h}$ , the estimation may give inconsistency, but the probability of occurrence of those  $\{\mathbf{h}\}$  which yield inconsistency is zero). For the MLE, define  $\mathbf{R}_1(\Delta) = (\mathbf{I} - \Sigma_B) \mathbf{U}^H \mathbf{\Gamma}(\Delta) \mathbf{U} \Sigma_S \mathbf{V}^H$ . Since  $\mathbf{h}$  is a continuous random vector, if  $\mathbf{R}_1(\Delta \neq 0) \neq \mathbf{0}_{N \times L}$ , the probability of  $(z_1(\Delta) = \mathbf{R}_1(\Delta) \mathbf{h} = \mathbf{0}_N)$  is equal to zero, i.e., the probability of  $(G_1(\Delta \neq 0) = 0)$  is zero. Hence, as long as  $\mathbf{R}_1(\Delta \neq 0) \neq \mathbf{0}_{N \times L}$ , the consistency in the probabilistic sense is achieved. This condition can be expressed more compactly as

$$\mathbf{U}_Z^H \mathbf{\Gamma}(\Delta) \mathbf{U}_P \neq \mathbf{0}_{(N-\mu) \times \mu} \quad \forall \Delta \neq 0 \quad (44)$$

where

$$\mathbf{U}_P = [\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{\mu-1}] \quad (45)$$

$$\mathbf{U}_Z = [\mathbf{u}_\mu, \mathbf{u}_{\mu+1}, \dots, \mathbf{u}_{N-1}]. \quad (46)$$

For a pilot tone vector  $\mathbf{c}$  with the nonzero pilot tone index set  $\{t_k : k = 1, 2, \dots, P\}$ , a sufficient pilot design condition for the above consistency in the probabilistic sense is given by

$$1 \leq P < \frac{(N+L)}{2} \& \{t_k^{(l)}\} \neq \{t_k\} \quad \forall l \in \{1, \dots, N-1\}. \quad (47)$$

The proof is given in Appendix C.

For the estimator in [13], define  $\mathbf{R}_2(\Delta) = \mathbf{F}_Z^T \mathbf{\Gamma}(\Delta) \mathbf{F}^H$ .  $\mathbf{Hc} = \mathbf{F}_Z^T \mathbf{\Gamma}(\Delta) \mathbf{F}^H \mathbf{C} \mathbf{F}_L$ . Then, similar to the MLE case, the

consistency in the probabilistic sense for the estimator of [13] is achieved if  $\mathbf{R}_2(\Delta \neq 0) \neq \mathbf{0}_{(N-P) \times L}$ . This condition can be compactly expressed as

$$\mathbf{F}_Z^T \mathbf{\Gamma}(\Delta) \mathbf{F}_P^* \neq \mathbf{0}_{(N-P) \times P}. \quad (48)$$

A sufficient pilot design condition for the above consistency in the probabilistic sense is given by

$$\{t_k^{(l)}\} \neq \{t_k\} \quad \forall l \in \{1, \dots, N-1\}. \quad (49)$$

The proof is given in Appendix C. Note that the pilot design used in [13] which does not satisfy  $P > L$  yields the consistency in the probabilistic sense only.

#### VI. CONSISTENT PILOT DESIGN FOR AN ARBITRARY SMALLER ESTIMATION RANGE

The above consistent pilot design addresses consistency over the entire estimation range  $|v| < N/2$ , which may be neither necessary nor economical for some consumer-related wireless systems. In this section, we develop generalized consistent pilot designs for an arbitrary but fixed estimation range  $|v| < M/2$  where  $M \leq N$ . These generalized designs include the consistent designs from Section IV, which also serve as building blocks, as a special case of  $M = N$ .

Based on (31) with  $|v| < M/2$  and hence  $\Delta \in (v - M/2, v + M/2)$ , we obtain the following sufficient pilot design condition for the consistency over the estimation range  $|v| < M/2$ .

*Sufficient Condition:*  $(N - P) \geq P$ , and for any cyclic-shifting distance  $l \in \{-\lfloor M/2 \rfloor, \dots, -1, 1, \dots, \lfloor M/2 \rfloor - 1\}$ ,  $\mathbf{c}^{(l)}$  always has at least  $L$  nonzero pilot tones located at the null-tone indices of the original unshifted pilot vector  $\mathbf{c}$ .

We can equivalently express the above condition as

$$(\mathbf{1}_N - \mathbf{p})^T \mathbf{p}^{(l)} \geq L, \text{ or } \phi(l) \leq P - L \\ \forall l \in \left\{ -\left\lfloor \frac{M}{2} \right\rfloor, \dots, -1, 1, \dots, \left\lfloor \frac{M}{2} \right\rfloor - 1 \right\} \quad (50)$$

where  $\mathbf{p}$  is the pilot-tone location vector whose elements are given by  $[\mathbf{p}]_n = \delta[n - t_k]$  and  $\phi(l) = \mathbf{p}^T \mathbf{p}^{(l)}$  is the *periodic autocorrelation* function of  $\mathbf{p}$ . Based on the above sufficient condition for  $|v| < M/2$ , we present several consistent design patterns in the following.

##### A. Design Pattern Group 1

Recall that for consistency, the number of pilot tones  $P$  must be greater than or equal to the number of channel taps  $L$ . Denote the pilot spacing between  $i$ th and  $((i+1) \bmod P)$ th pilots as  $d_i$ , i.e.,

$$d_i = (t_{(i+1) \bmod P} - t_i) \bmod N, \quad i = 0, 1, \dots, P-1. \quad (51)$$

If we design

$$d_i \geq M \quad \forall i \in \{0, 1, \dots, P-1\} \quad (52)$$

then (50) is satisfied. Hence, this condition in (52) guarantees consistency over the range  $|v| < M/2$ . Note that this design includes periodic training signals.

### B. Design Pattern Group 2

We construct  $N \times 1$  pilot tone vectors from smaller size  $M_i \times 1$  subvectors  $\{\mathbf{c}_i\}$ , whose corresponding pilot location subvectors  $\{\mathbf{q}_i\}$  are defined by

$$[\mathbf{q}_i]_k = \begin{cases} 1, & \text{if } [\mathbf{c}_i]_k \neq 0 \\ 0, & \text{otherwise} \end{cases} \quad (53)$$

and

$$(\mathbf{1}_{M_i} - \mathbf{q}_i)^T \mathbf{q}_i^{(l)} \geq L_i \\ \forall l = -\left\lfloor \frac{M}{2} \right\rfloor, \dots, -1, 1, \dots, \left\lfloor \frac{M}{2} \right\rfloor - 1 \quad (54)$$

with  $M_i \geq M$ . Our designs from Section IV can be used to obtain  $\{\mathbf{c}_i\}$ . We can separate  $\mathbf{c}_i$  into two parts as  $\mathbf{c}_i = [\mathbf{c}_{i-}^T, \mathbf{c}_{i+}^T]^T$ , where  $\mathbf{c}_{i-}^T$  is the  $\lceil (M_i/2) \rceil \times 1$  vector with corresponding pilot location subvector  $\mathbf{q}_{i-}^T$ , and  $\mathbf{q}_i = [\mathbf{q}_{i-}^T, \mathbf{q}_{i+}^T]^T$ . Under the above setup, we develop several designs that are consistent over  $|v| < M/2$  in the following.

1) *Design Pattern Group 2-A*: Design the pilot vector  $\mathbf{c}$  to contain  $n_1$  cyclically consecutive repetitions of  $\mathbf{c}_1$ . For  $M_1 n_1 + M \leq N$ , insert  $\lfloor (M/2) \rfloor$  and  $\lfloor (M/2) \rfloor - 1$  null tones at the left and right ends of the whole repetition, respectively, i.e., the subvector of  $\mathbf{c}$  is  $[\mathbf{0}_{\lfloor (M/2) \rfloor}^T, \mathbf{c}_1^T, \dots, \mathbf{c}_1^T, \mathbf{0}_{\lfloor (M/2) \rfloor - 1}^T]^T$ , and the remaining tones can be arbitrary. For  $N - M < M_1 n_1 \leq N$ , the remaining  $N - M_1 n_1$  tones are set as null tones. Then if we design

$$n_1 L_1 \geq L \quad (55)$$

(50) is satisfied, and hence, we obtain consistency over the range  $|v| < M/2$ .

2) *Design Pattern Group 2-B*: If the two null-tone blocks mentioned in the design pattern group 2-A are excluded, then the consistent condition is changed to

$$(n_1 - 1)L_1 + \bar{L}_1 \geq L \quad (56)$$

where

$$\bar{L}_1 = \min(\beta_1^+, \beta_1^-) \\ \beta_1^+ = \min\{[q_1(0), \dots, q_1(M_1 - 1 - l)] \\ \times (\mathbf{1}_{(M_1-l)} - [q_1(l), \dots, q_1(M_1 - 1)]^T)\}, \quad (57)$$

$$l = 1, \dots, \left\lfloor \frac{M}{2} \right\rfloor - 1 \quad (58)$$

$$\beta_1^- = \min\{[q_1(l), \dots, q_1(M_1 - 1)] \\ \times (\mathbf{1}_{(M_1-l)} - [q_1(0), \dots, q_1(M_1 - l - 1)]^T)\}, \\ l = -\left\lfloor \frac{M}{2} \right\rfloor, \dots, -1. \quad (59)$$

For  $n_1 > 1$ , neglecting  $\bar{L}_1$  in (56) gives a simpler sufficient design condition as

$$(n_1 - 1)L_1 \geq L. \quad (60)$$

3) *Design Pattern Group 2-C*: Design  $\mathbf{c}$  to contain  $n$  subvectors  $\{[\mathbf{c}_{i-}^T, \mathbf{c}_{i+}^T, \mathbf{c}_{i-}^T]^T : i = 1, \dots, n\}$ , and  $\sum_{i=1}^n (M_i + \lceil (M_i/2) \rceil) \leq N$ . These subvectors need not be consecutively located. Then the condition

$$\sum_{i=1}^n (L_i + \bar{L}_i) \geq L \quad (61)$$

satisfies (50), and hence, provides consistency over the range  $|v| < M/2$ , where we have (62)–(64) shown at the bottom of the page. For  $n > 1$ , neglecting  $\bar{L}_i$  in (61), we obtain a simpler sufficient condition as

$$\sum_{i=1}^n L_i \geq L. \quad (65)$$

## VII. ROBUST CONSISTENT PILOT DESIGNS

From the simulation results of the CFO estimation of the consistent pilots (the results will be presented in Section VIII), we observe that the CFO estimation performance gaps between different consistent pilot designs can be quite significant at moderate-to-low SNR range (5–15 dB for most circumstances). Below and above this range, no obvious difference is observed. These facts point out that the outlier statistics of consistent pilot designs can be significantly different from one another.

The consistent condition in (31) is derived from the following noise-free version of the cost function in (18):

$$\bar{g}(\hat{v}) = \mathbf{h}^H \mathbf{Q}(\Delta) \mathbf{h} \quad (66)$$

$$\bar{L}_i = \min(\beta_i^+, \beta_i^-) \quad (62)$$

$$\beta_i^+ = \min\left\{ [q_{i-}(0), \dots, q_{i-}(\left\lfloor \frac{M_i}{2} \right\rfloor - 1 - l)] \left( \mathbf{1}_{\lceil (M_i/2) \rceil - l} - [q_{i-}(l), \dots, q_{i-}(\left\lfloor \frac{M_i}{2} \right\rfloor - 1)]^T \right) \right\}, \\ l = 1, \dots, \left\lfloor \frac{M}{2} \right\rfloor - 1 \quad (63)$$

$$\beta_i^- = \min\left\{ [q_{i-}(l), \dots, q_{i-}(\left\lfloor \frac{M_i}{2} \right\rfloor - 1)] \left( \mathbf{1}_{\lceil (M_i/2) \rceil - l} - [q_{i-}(0), \dots, q_{i-}(\left\lfloor \frac{M_i}{2} \right\rfloor - l - 1)]^T \right) \right\}, \\ l = -\left\lfloor \frac{M}{2} \right\rfloor, \dots, -1 \quad (64)$$

where

$$\mathbf{Q}(\Delta) = \mathbf{S}^H \mathbf{\Gamma}(\Delta) \mathbf{B} \mathbf{\Gamma}^H(\Delta) \mathbf{S} \quad (67)$$

$$= \mathbf{U}_q \text{diag}\{\lambda_1(\Delta), \lambda_2(\Delta), \dots, \lambda_L(\Delta)\} \mathbf{U}_q^H \quad (68)$$

and  $\{\lambda_1(\Delta), \lambda_2(\Delta), \dots, \lambda_L(\Delta)\}$  are the ordered eigenvalues of  $\mathbf{Q}(\Delta)$  such that  $\lambda_i(\Delta) \geq \lambda_{i+1}(\Delta)$ , and  $\mathbf{U}_q$  is a unitary matrix with the  $i$ th column  $\mathbf{u}_{qi}$  representing the eigenvector of  $\mathbf{Q}(\Delta)$  corresponding to  $\lambda_i(\Delta)$ .

If  $\bar{g}(\tilde{v})$  fluctuates more drastically with the estimation trial value  $\tilde{v} (\neq v)$ , the estimation will be less robust to noise or more prone to outliers. The characterization of the fluctuation of  $\bar{g}(\tilde{v})$  reflects how likely the other noise-free local maxima are to become close to the noise-free global maximum value of  $\bar{g}(\tilde{v})$ . If some noise-free local maxima are quite close to the noise-free global maximum value, then in the presence of a moderate-to-high noise level (relative to the snapshot received signal level), any one of them could be often selected by the estimator, yielding outlier occurrence.

In order to suppress the occurrence of the outliers, we need to reduce the fluctuation of the cost function  $\bar{g}(\tilde{v})$ . It is obvious that  $\mathbf{Q}(\Delta)$  depends on  $\Delta$  and  $\mathbf{S}$ , and if  $\mathbf{h} = a_i \mathbf{u}_{qi}$  where  $a_i$  is a scalar, then  $\bar{g}(\tilde{v}) = |a_i|^2 \lambda_i(\Delta)$ . For a fixed snapshot channel energy, the maximum and minimum values of  $\bar{g}(\tilde{v})$  are  $\lambda_1(\Delta)$  and  $\lambda_L(\Delta)$  times the snapshot channel energy, respectively, i.e.,

$$\max \left( \frac{\bar{g}(\tilde{v})}{\mathbf{h}^H \mathbf{h}} \right) = \lambda_1(\Delta) \quad (69)$$

$$\min \left( \frac{\bar{g}(\tilde{v})}{\mathbf{h}^H \mathbf{h}} \right) = \lambda_L(\Delta). \quad (70)$$

We observe from the above equations that the spread of the eigenvalues  $\{\lambda_i\}$  indicates the fluctuation of the cost function over different  $\mathbf{h}$ . Thus, our design for robustness against outliers becomes finding training signals that give the smallest eigenspread of  $\mathbf{Q}(\Delta)$ . We measure the eigenspread in two ways: the ratio of maximum to minimum eigenvalues (max-min-ratio) and the product of all  $L$  eigenvalues (the determinant). A smaller max-min-ratio or a larger determinant (due to the arithmetic and geometric means inequality) implies a smaller eigenspread. Hence, our robustness design criteria become

$$\min \left( C_1 \triangleq \sum_{n=1}^{N_s} \frac{\lambda_1(\Delta_n)}{\lambda_L(\Delta_n)} \right) \quad (71)$$

$$\max \left( C_2 \triangleq \sum_{n=1}^{N_s} \prod_{i=1}^L \lambda_i(\Delta_n) \right) \quad (72)$$

where we have evaluated for the whole CFO estimation range using  $N_s = KN$  ( $K$  is the up-sampling factor) equispaced trial points  $\{\Delta_n\}$ .

The above criteria need eigendecompositions of  $N_s$  matrices due to the dependence of  $\mathbf{Q}(\Delta)$  on  $\Delta$ . We obtain a simpler version by removing the dependency on  $\Delta$  or equivalently replacing  $\mathbf{\Gamma}(\Delta)$  with an identity matrix in (67). Then the simplified version of  $\mathbf{Q}(\Delta)$ , denoted by  $\bar{\mathbf{Q}}$ , becomes

$$\bar{\mathbf{Q}} = \mathbf{S}^H \mathbf{S}. \quad (73)$$

Denote the ordered eigenvalues of  $\bar{\mathbf{Q}}$  as  $\delta_1, \dots, \delta_L$ , where  $\delta_i \geq \delta_{i+1}$ . Then we obtain the following simplified robustness design criteria:

$$\min \left( \bar{C}_1 \triangleq \frac{\delta_1}{\delta_L} \right) \quad (74)$$

$$\max \left( \bar{C}_2 \triangleq \prod_{i=1}^L \delta_i \right). \quad (75)$$

All the above robustness design criteria when applied to the consistent pilots yield essentially the same robust and consistent pilot patterns, as will be shown in Section VIII.

### VIII. SIMULATION RESULTS AND DISCUSSIONS

We consider an OFDM system with  $N = 128$  subcarriers in a multipath Rayleigh fading channel with  $L = 8$  (unless otherwise stated) sample-spaced taps having an exponential power delay profile with a 3-dB-per-tap decaying factor. Simulation results are obtained from  $10^4$  independent runs. Two estimators denoted by ‘‘MLE’’ from [14] and ‘‘JL’’ (Jing Lei) from [13] are considered. We use the following four different preambles (of one OFDM symbol duration each): 1) the distinctively spaced pilot tones with  $P < L$  used in [13] (denoted by ‘‘JL’s pilot’’); 2) a repetitive preamble consisting of 16 identical parts (excluding the CP) generated by 16 pilot tones located at the subcarrier indices  $\{8i : i = 0, \dots, 15\}$  (denoted by ‘‘repetitive’’); 3) the proposed consistent pilot tones based on an m-sequence (denoted by ‘‘proposed (m-seq.)’’ with  $P = 16$ ; and 4) the proposed consistent pilot tones based on distinctive spacing with  $P = 9 > L$  (denoted by ‘‘proposed (distinctive)’’).

In Fig. 1, we compare the MSEs of the two estimators. If  $P > L$  (proposed (m-seq.)), the MLE gives a slightly better MSE, which corroborates our proof in Appendix A. Note that the MSE difference is not significant, and hence, JL’s estimator would be a good choice due to its complexity advantage. Our purpose here is just to point out in a strictly theoretical sense that the two estimators are not the same. The theoretical (quasi-analytical) MSEs<sup>5</sup> are also included and show a good agreement with the simulation results, except at low SNR due to the outliers. In practice, the outlier occurrence will be much less frequent since the signal will not be detected if the channel is in fade (CFO estimation will only be performed after the (sync) preamble is detected).

Fig. 2 presents MSEs obtained with different pilot signals. Since all pilot signals used satisfy the condition for estimation consistency in the probabilistic sense, no significant MSE differences are observed in the simulation results<sup>6</sup> (i.e.,  $\{\mathbf{h}\}$  which yield inconsistency for the not-absolutely-consistent pilot signals do not occur in the  $10^4$  channel realizations).

In Figs. 3–6, we illustrate the consistency/inconsistency of several pilot signals. We plot the MLE metrics (cost functions

<sup>5</sup>The theoretical MSE given in Appendix A for a particular  $\mathbf{h}$  is averaged over the random channel by simulation.

<sup>6</sup>Note that the MSE simulation results at low SNR are not as reliable as those at high SNR, due to different occurrences of outliers for different preambles (different preambles’ metrics change differently in response to different  $\mathbf{h}$ , and the number of outlier occurrences within the  $10^4$  simulation runs will not be the same for different preambles).

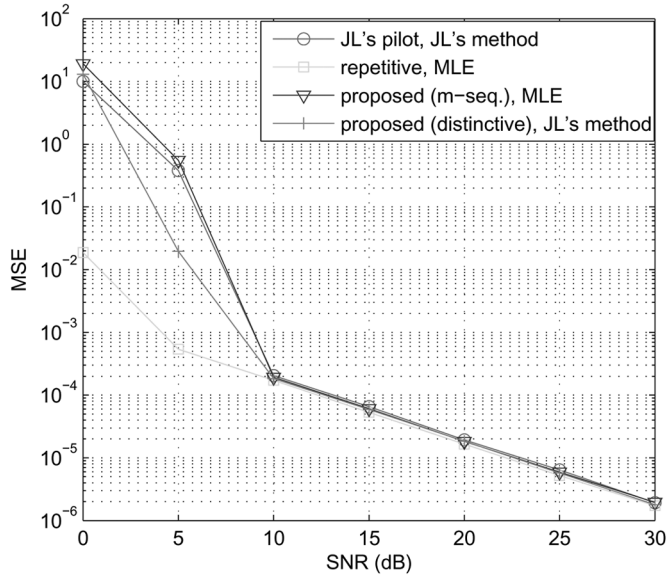


Fig. 2. Simulated MSE performance of different preambles.

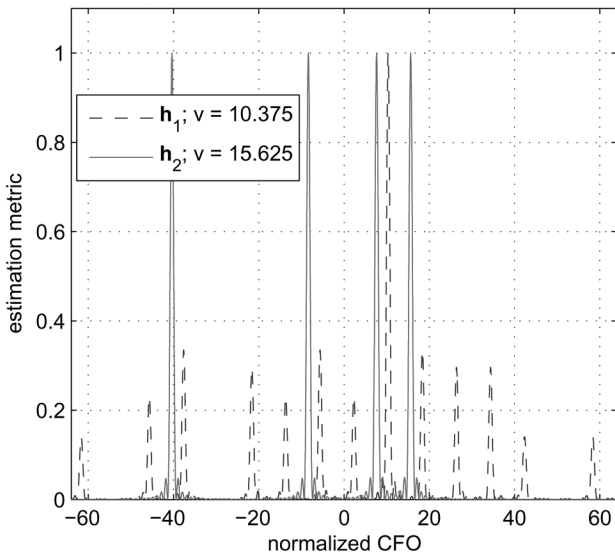


Fig. 3. Noise-free CFO estimation metrics (normalized) of JL's pilot signal for two channel realizations.

to maximize) of each of the pilot signals for two channel realizations (denoted by  $\mathbf{h}_1$  and  $\mathbf{h}_2$  in the figures) in the absence of noise. We use different CFOs for the two channel realizations for the convenience of presenting two metrics on the same plot. The metric for JL's pilot is consistent for the first channel realization, but inconsistent for the second channel realization, due to the numerous metric maxima. The repetitive pilot signal metric has several periodic maxima which limit its estimation range. Note that the metric periodicity of the repetitive preamble is still preserved in the presence of noise. This fact ensures the consistency (within a limited range) for the repetitive preamble if the CFO is within its limited estimation range. The proposed consistent pilot signals yield the metrics with a single maximum for all channel realizations.

Fig. 7 depicts the MSEs of different pilot signals for a fixed channel which yields inconsistency for JL's pilot signal. Due to

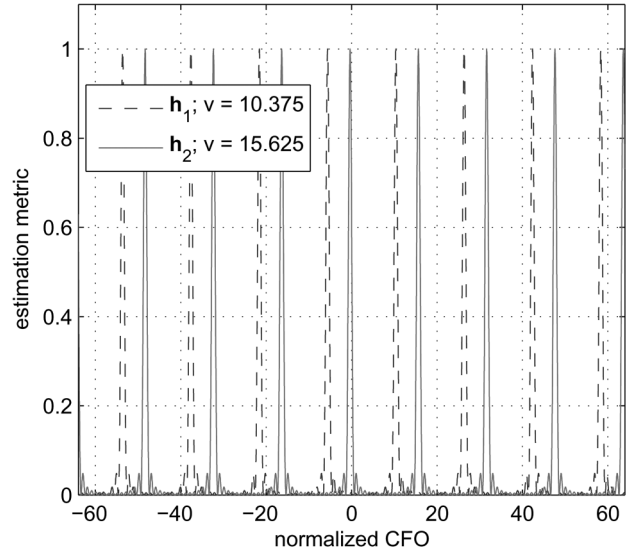


Fig. 4. Noise-free CFO estimation metrics (normalized) of the repetitive preamble for two channel realizations.

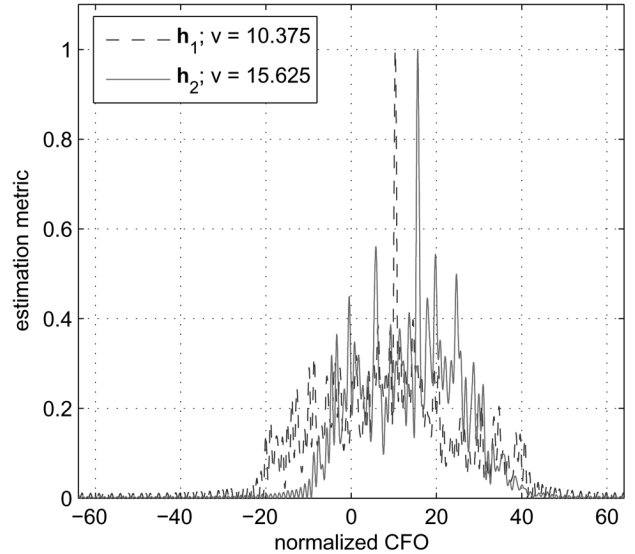


Fig. 5. Noise-free CFO estimation metrics (normalized) of the proposed m-sequence-based pilot signal for two channel realizations.

this inconsistency, a complete estimation failure is observed for JL's pilot signal. Fig. 8 shows the estimation ranges of different pilot signals. The repetitive pilot signal has a limited estimation range of  $|v| < 8$ , while the others can handle the maximum range of  $|v| < N/2 = 64$ .

Next, we investigate the consistency of the preambles (for coarse CFO estimation) used in IEEE 802.11a ( $N = 64$ ) and 802.16a ( $N = 256$ ) OFDM systems. We find that the IEEE 802.11a preamble (of length  $N = 64$  generated from (6) [1, p. 12]) does not yield CFO estimation consistency for channels with  $L > 8$ . Similarly, the IEEE 802.16a preamble (of length  $N = 256$  generated from (78) [3, p. 447]) does not yield CFO estimation consistency for channels with  $L > 26$ . For both systems, the channel lengths can easily be in the corresponding inconsistent ranges mentioned above. The MLE metrics associated with the IEEE 802.11a and 802.16a OFDM preambles are

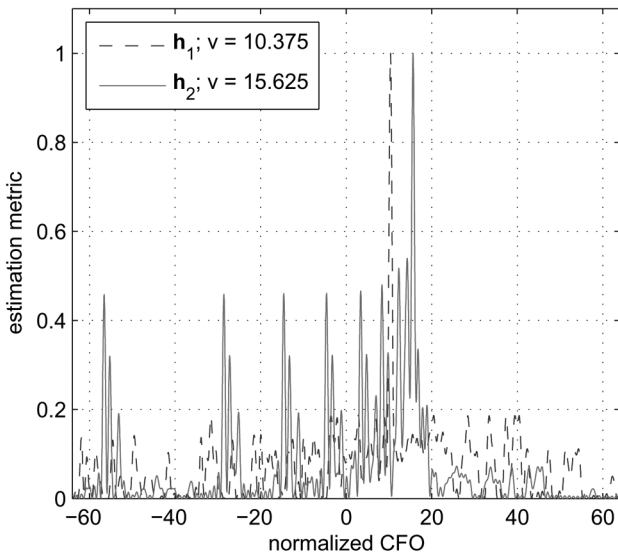


Fig. 6. Noise-free CFO estimation metrics (normalized) of the proposed distinctive-spacing-based pilot signal for two channel realizations.

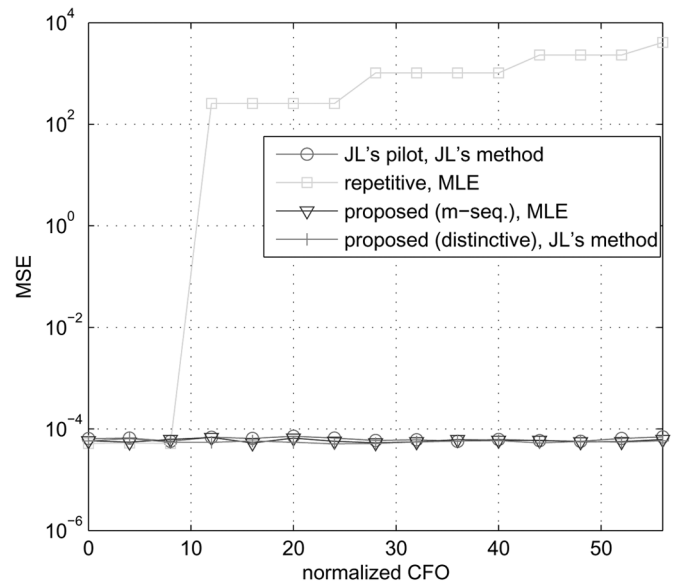


Fig. 8. CFO estimation ranges associated with different preambles.

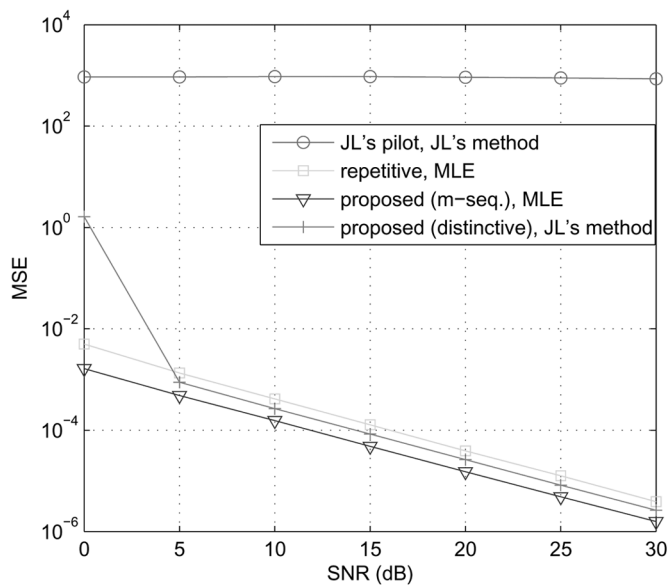


Fig. 7. MSE performance associated with different preambles for a particular channel realization which yields inconsistency for JL's pilot signal.

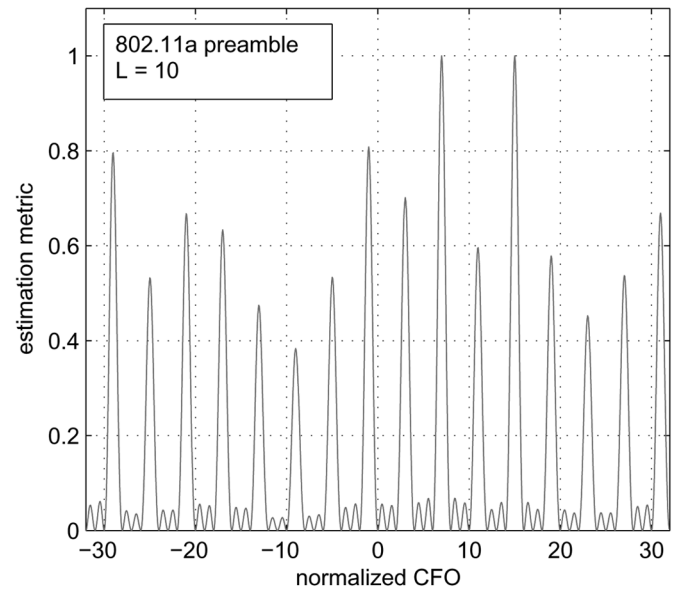


Fig. 9. CFO estimation metrics (normalized) associated with the IEEE 802.11a OFDM preamble.

presented in Figs. 9 and 10, respectively, for some channel realizations. The estimation inconsistencies for both preambles are observed, as can be seen from the two maxima in each metric.

We illustrate the reliability of our robustness design criteria in Figs. 11–13, where we use all possible pilots obtained from the designs of distinctive spacing, m-sequence, and BCH codes as described in Section IV.

In the top figures of Figs. 11–13, we present our robustness design metrics  $C_1$  and  $\bar{C}_2$  [from (71) and (75)] of all consistent pilots mentioned above for different system and channel parameters. The middle figures show two independent sets (denoted by solid and dashed lines, respectively) of the corresponding simulation MSE results. Since there are no noticeable performance differences among all consistent pilots at very low or high SNR,

we only present results for SNRs of 5, 10, and 15 dB. The trajectories of  $C_i$  and  $\bar{C}_i$  are almost the same, and hence, we only present  $C_1$  and  $\bar{C}_2$ .

All proposed robustness criteria give reliable results. Large differences in robustness metric values result in large MSE gaps. For comparable robustness metric values, the corresponding MSE results may not show the same trends of the robustness metric values. This can be explained by the fact that a pilot signal with a smaller probability of outlier occurrence, but much larger estimation errors in these outliers than another pilot signal, may give a larger MSE. On the other hand, for two CFO estimation errors larger than a certain threshold, there would be no bit-error rate (BER) difference regardless of the value of the CFO estimation error difference. Hence, MSE

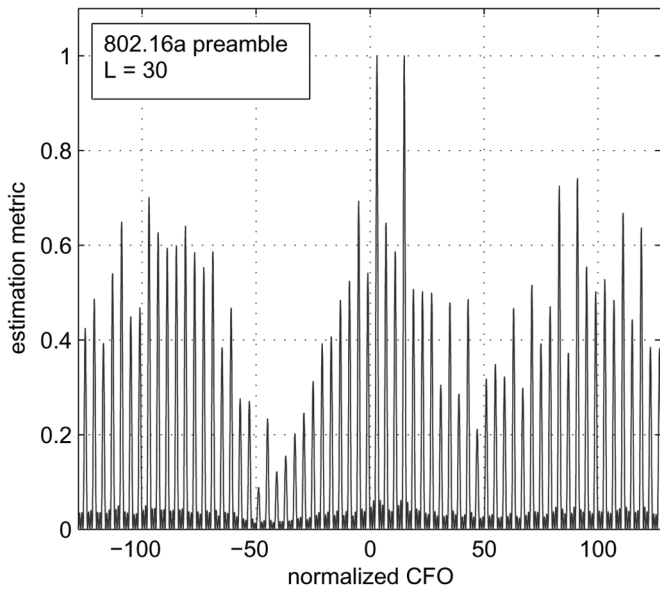


Fig. 10. CFO estimation metrics (normalized) associated with the IEEE 802.16a OFDM preamble.

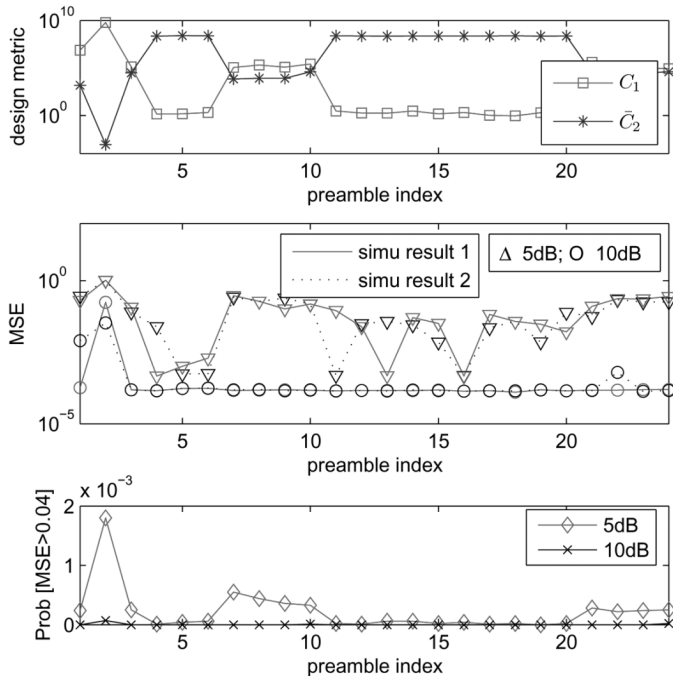


Fig. 11. Proposed robustness design metric (top), the simulation MSE performance (middle), and the probability of outlier (defined by  $\text{MSE} > 0.04$ ) (bottom) for different consistent pilot designs, with  $N = 128$  and  $L = 8$ .

results do not completely characterize the robustness against outliers.

A better illustration of robustness against outliers is to show the probability of estimation error (or MSE) larger than a certain threshold. We present this outlier probability for different system and channel parameters in the bottom figures of Figs. 11–13. The distinctive spacing, the  $m$ -sequence, and the BCH-code-based designs correspond to the indices  $\{1\}$ ,  $\{2-6\}$ ,  $\{7-24\}$  in Fig. 11,  $\{1-5\}$ ,  $\{6-10\}$ ,  $\{11-40\}$  in Fig. 12, and

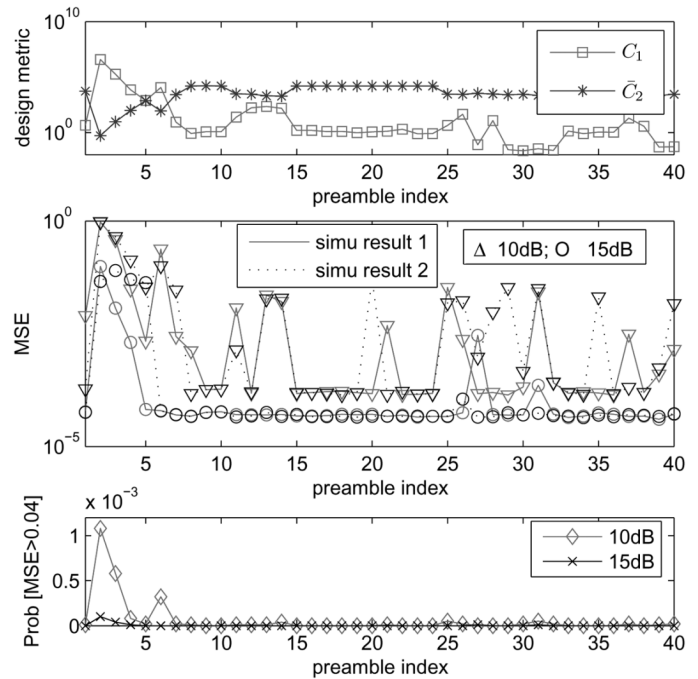


Fig. 12. Proposed robustness design metric (top), the simulation MSE performance (middle), and the probability of outlier (defined by  $\text{MSE} > 0.04$ ) (bottom) for different consistent pilot designs, with  $N = 128$  and  $L = 4$ .

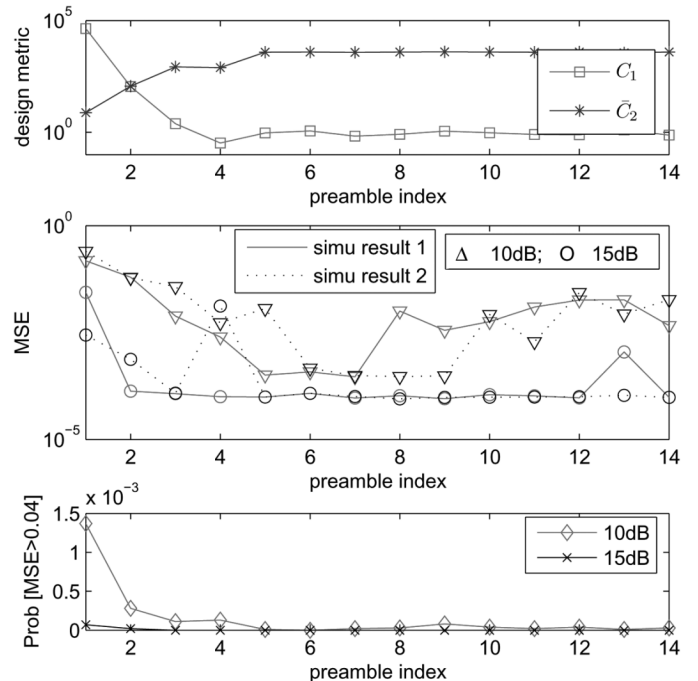


Fig. 13. Proposed robustness design metric (top), the simulation MSE performance (middle), and the probability of outlier (defined by  $\text{MSE} > 0.04$ ) (bottom) for different consistent pilot designs, with  $N = 64$  and  $L = 4$ .

$\{1-3\}$ ,  $\{4-6\}$ ,  $\{7-14\}$  in Fig. 13, respectively. Which design gives the most robust performance depends on the system and channel parameters. But we notice from Figs. 11–13 that the distinctive spacing designs most often do not yield robust performance.

## IX. CONCLUSIONS

The existing OFDM preambles or pilot signals may not guarantee an absolute consistency of the CFO estimation, and can result in a complete failure in frequency synchronization (and hence, no meaningful communication) for certain channel conditions. For emergency and other critical communications systems, absolute consistency should not be compromised, and our proposed pilot designs guarantee the absolute consistency over the considered estimation range (either the maximum range of the half-sampling rate or a smaller range). Our CFO estimation consistency criteria can also be used to determine the maximum channel length for which a particular preamble will yield CFO estimation consistency. Furthermore, the proposed robustness design criteria enable us to choose the consistent pilots with better MSE performance at moderate or low SNR or greater robustness against outliers. In summary, our proposed robust and consistent pilot designs provide wireless links with less outage and better resilience.

## APPENDIX A

In this appendix, we will prove that if  $P > L$ , the MSE of the MLE in [14] is smaller than that of the estimator in [13], and if  $P \leq L$ , they give the same MSE performance. For a given  $\mathbf{h}$ , the MSEs of the MLE and JL methods, denoted by  $\text{MSE}_1$  and  $\text{MSE}_2$ , respectively, are given by (see [13], [14])

$$\text{MSE}_i = \frac{\sigma_n^2}{2\mathbf{y}^H(\mathbf{I} - \mathbf{B}_i)\mathbf{y}}, \quad i = 1, 2 \quad (76)$$

where

$$\mathbf{y} = \frac{2\pi\mathbf{M}\mathbf{S}\mathbf{h}}{N} \quad (77)$$

$$\mathbf{M} = \text{diag}\{0, 1, \dots, N-1\}. \quad (78)$$

From (76), it is clear that we just need to show that  $\mathbf{y}^H(\mathbf{B}_2 - \mathbf{B}_1)\mathbf{y} \geq 0$  in order to prove that  $\text{MSE}_1 \leq \text{MSE}_2$ .

Note that  $\{\mathbf{B}_i\}$  are projection matrices (i.e.,  $\mathbf{B}_i^2 = \mathbf{B}_i$ ). Using  $\mathbf{B}_2 = \mathbf{F}^H \text{diag}\{\mathbf{p}\}\mathbf{F}$  and (15), we also have

$$\mathbf{B}_2\mathbf{B}_1 = \mathbf{B}_1\mathbf{B}_2 = \mathbf{B}_1 \quad (79)$$

$$(\mathbf{B}_2 - \mathbf{B}_1)^2 = \mathbf{B}_2^2 - \mathbf{B}_2\mathbf{B}_1 - \mathbf{B}_1\mathbf{B}_2 + \mathbf{B}_1^2 = \mathbf{B}_2 - \mathbf{B}_1 \quad (80)$$

from which we can see that  $(\mathbf{B}_2 - \mathbf{B}_1)$  is also a projection matrix. The eigenvalues of projection matrices are either ones or zeros, and hence, we obtain

$$\text{rank}(\mathbf{B}_2 - \mathbf{B}_1) = \text{trace}(\mathbf{B}_2 - \mathbf{B}_1) = P - \min(P, L) \quad (81)$$

where we have used  $\text{rank}(\mathbf{B}_1) = \min(P, L)$  for which a proof will be provided shortly.

If  $P > L$ ,  $(\mathbf{B}_2 - \mathbf{B}_1)$  is a positive-definite matrix, and  $\mathbf{y}^H(\mathbf{B}_2 - \mathbf{B}_1)\mathbf{y} > 0$ ; if  $P \leq L$ ,  $(\mathbf{B}_2 - \mathbf{B}_1) = \mathbf{0}$  and  $\mathbf{y}^H(\mathbf{B}_2 - \mathbf{B}_1)\mathbf{y} = 0$ . From the above result, we immediately see that if  $P \leq L$ ,  $\text{MSE}_1 = \text{MSE}_2$  and otherwise,  $\text{MSE}_1 < \text{MSE}_2$ . This result holds for any channel realization  $\mathbf{h}$  ( $\neq \mathbf{0}_L$ ), and hence, for the random fading channel (for which the  $(\text{MSE}_i)$  are averaged over the random channel realizations) as well.

A proof for  $\text{rank}(\mathbf{B}_1) = \min(P, L)$  is as follows. Equation (15) can be expressed as

$$\begin{aligned} \mathbf{S} &= \sqrt{N}\mathbf{F}^H\boldsymbol{\Theta}\boldsymbol{\Theta}^T\mathbf{C}\boldsymbol{\Theta}\boldsymbol{\Theta}^T\mathbf{F}_L \\ &= \begin{cases} \sqrt{N}\tilde{\mathbf{F}}_N^H\tilde{\mathbf{C}}_{N \times L}\tilde{\mathbf{F}}_L, & \text{if } P \leq L \\ \sqrt{N}\tilde{\mathbf{F}}_N^H\tilde{\mathbf{C}}_{N \times P}\tilde{\mathbf{F}}_P [\mathbf{I}_L, \mathbf{0}_{L \times (P-L)}]^T, & \text{if } P > L \end{cases} \end{aligned} \quad (82)$$

where

$$\tilde{\mathbf{F}}_k = [\mathbf{I}_k, \mathbf{0}_{k \times (N-k)}]\boldsymbol{\Theta}^T\mathbf{F}[\mathbf{I}_k, \mathbf{0}_{k \times (N-k)}]^T \quad (83)$$

$$\tilde{\mathbf{C}}_{N \times k} = \boldsymbol{\Theta}^T\mathbf{C}\boldsymbol{\Theta}[\mathbf{I}_k, \mathbf{0}_{k \times (N-k)}]^T. \quad (84)$$

Note that  $\{\tilde{\mathbf{F}}_k\}$  are Vandermonde matrices with all distinct rows and all distinct columns, and hence, they are full-rank matrices. Then, from (82), we immediately see that  $\text{rank}(\mathbf{S}) = \min(P, L)$  and hence,  $\text{rank}(\mathbf{B}_1) = \min(P, L)$  [see (26) and (27)].

## APPENDIX B

In this appendix, we will prove that the proposed pilot design sufficient condition given in Section IV satisfies the consistency condition in (31) for the MLE and in (37) for JL's estimator. Note that  $P > L$  in our design condition.

First, we will prove it for the MLE. Express the orthogonal vectors  $\{\mathbf{u}_1, \dots, \mathbf{u}_L\}$  in (32) as linear transforms of the Fourier orthogonal vectors as

$$\mathbf{U}_1 = \mathbf{F}_P^*\mathbf{D}. \quad (85)$$

Define

$$\mathbf{D} = \mathbf{F}_P^T\mathbf{U}_1 \quad (86)$$

$$\mathbf{E} = \mathbf{V}\boldsymbol{\Sigma}_S^H(\boldsymbol{\Sigma}_S\boldsymbol{\Sigma}_S^H)^{-1}[\mathbf{I}_L, \mathbf{0}_{L \times (N-L)}]^T. \quad (87)$$

Then,  $\mathbf{U}_1$  in (32) can be expressed as

$$\mathbf{U}_1 = \mathbf{U}\boldsymbol{\Sigma}_S\mathbf{V}^H\mathbf{E} = \sqrt{N}\mathbf{F}^H\mathbf{C}\mathbf{F}_L\mathbf{E} \quad (88)$$

and (86) becomes

$$\mathbf{D} = \sqrt{N}\mathbf{F}_P^T\mathbf{F}^H\mathbf{C}\mathbf{F}_L\mathbf{E} = \sqrt{N}[\mathbf{I}_P, \mathbf{0}_{P \times (N-P)}]\boldsymbol{\Theta}^T\mathbf{C}\mathbf{F}_L\mathbf{E}. \quad (89)$$

Note that  $\mathbf{E}$  is an  $L \times L$  full-rank matrix and any  $K (\geq L)$  rows of  $\mathbf{F}_L$  form a  $K \times L$  Vandermonde matrix with all distinct rows and distinct columns, i.e., a rank- $L$  matrix. The left-multiplication by  $\boldsymbol{\Theta}^T$  shifts rows at indices  $\{t_k\}$  to indices  $\{0, 1, \dots, P-1\}$ . Hence, we immediately obtain that a matrix formed by any  $K (\geq L)$  rows of  $\mathbf{D}$ , denoted by  $\mathbf{D}_K$ , is a full-rank matrix, i.e.,  $\text{rank}(\mathbf{D}_K) = L$ .

Next,  $\mathbf{U}_2^H$  can be decomposed as

$$\mathbf{U}_2^H = \mathbf{A}[\mathbf{F}_Z, \mathbf{Q}]^T \quad (90)$$

where  $\mathbf{Q}$  is an  $N \times (P-L)$  matrix whose columns form the basis of the null space of the columns of  $\mathbf{U}_1$  within the subspace

of the columns of  $\mathbf{F}_P$ , and  $\mathbf{A}$  is an  $(N-L) \times (N-L)$  full-rank matrix. Then, from (86) and (90), we have

$$\mathbf{U}_2^H \mathbf{\Gamma}(\Delta) \mathbf{U}_1 = \mathbf{A} \begin{bmatrix} \mathbf{T}_1 \mathbf{D} \\ \mathbf{T}_2 \mathbf{D} \end{bmatrix} \quad (91)$$

where

$$\mathbf{T}_1 = \mathbf{F}_2^T \mathbf{\Gamma}(\Delta) \mathbf{F}_P^* \quad (92)$$

$$\mathbf{T}_2 = \mathbf{Q}^T \mathbf{\Gamma}(\Delta) \mathbf{F}_P^* \quad (93)$$

Consider the following two cases.

- 1) When  $\Delta$  is any nonzero integer, from (92), we have

$$\mathbf{T}_1 = [\mathbf{e}_{l_1}, \mathbf{e}_{l_2}, \dots, \mathbf{e}_{l_K}, \mathbf{0}_{P \times (P-K)}] \mathcal{I}_{P \times P} \quad (94)$$

$$\mathbf{T}_1 \mathbf{D} = \mathbf{D}_K \quad (95)$$

where  $l_k \in \{0, 1, 2, \dots, N-1\}$  and  $l_i \neq l_k$  if  $i \neq k$ .  $\mathbf{e}_{l_i}$  is the  $l_i$ th column of the  $(N-P) \times (N-P)$  identity matrix, and  $\mathcal{I}_{P \times P}$  denotes a  $P \times P$  permutation matrix determined by  $\Delta$ .

From our design condition that after cyclic-shifting, at least  $L$  pilots fall into the initial null tone locations, we have  $L \leq K \leq P$ . We have already proved that  $\text{rank}(\mathbf{D}_K) = L$ . Since  $\mathbf{A}$  is a full-rank square matrix, from (91), we immediately obtain that  $\text{rank}(\mathbf{U}_2^H \mathbf{\Gamma}(\Delta) \mathbf{U}_1) = L$ .

- 2) When  $\Delta$  is not an integer, from (92), we have

$$[\mathbf{T}_1]_{m,n} = \frac{(1 - e^{j2\pi\Delta}) \cdot e^{j2\pi n k_m / N}}{e^{j2\pi n k_m / N} - e^{j2\pi (t_{k_n} + \Delta) / N}} \triangleq \frac{c \cdot a_m}{a_m - b_n} \quad (96)$$

where  $c \triangleq (1 - e^{j2\pi\Delta})$ ,  $a_m \triangleq e^{j2\pi n k_m / N}$ ,  $b_n \triangleq e^{j2\pi (t_{k_n} + \Delta) / N}$ ,  $c \neq 0$ ,  $\{a_m \neq b_n \forall m, n\}$ ,  $\{a_m \neq a_n \forall m \neq n\}$ , and  $\{b_m \neq b_n \forall m \neq n\}$  for  $m = 1, 2, \dots, N-P$ ;  $n = 1, 2, \dots, P$ . Denote the square matrix formed by any  $P$  rows of  $\mathbf{T}_1$  by  $\bar{\mathbf{T}}_1$  (recall that  $(N-P) \geq P$  in our design condition). Then

$$\det(\bar{\mathbf{T}}_1) = c^P \prod_{m=2}^P \prod_{n=1}^{(m-1)} \frac{(b_m - b_n)(a_n - a_m)}{(a_m - b_n)(a_n - b_m)} \cdot \prod_{l=1}^P \frac{1}{a_l - b_l} \neq 0 \quad (97)$$

which shows that  $\mathbf{T}_1$  is a full-rank (tall) matrix, i.e.,  $\text{rank}(\mathbf{T}_1) = P$ . Define an  $(N-P) \times (N-P)$  full-rank matrix  $\tilde{\mathbf{T}}_1 = [\mathbf{T}_1, \mathbf{T}_1^\perp]$  and an  $(N-P) \times L$  matrix  $\tilde{\mathbf{D}} = [\mathbf{D}^T, \mathbf{0}_{L \times (N-2P)}]^T$ . Then, we have  $\mathbf{T}_1 \mathbf{D} = \tilde{\mathbf{T}}_1 \tilde{\mathbf{D}}$ , and hence,  $\text{rank}(\mathbf{T}_1 \mathbf{D}) = \text{rank}(\tilde{\mathbf{D}}) = L$  [21]. Multiplying by the square full-rank matrix  $\mathbf{A}$  does not change the rank. Hence,  $\text{rank}(\mathbf{U}_2^H \mathbf{\Gamma}(\Delta) \mathbf{U}_1) = L$  for any noninteger  $\Delta$ . From 1) and 2), we have established that the proposed pilot design condition is sufficient for the consistency condition in (31).

Finally, we will prove that our design condition satisfies the consistency condition in (37) for JL's estimator. Consider the following two cases.

- 3) If  $\Delta$  is an integer, we have already proved that  $\text{rank}(\mathbf{F}_2^T \mathbf{\Gamma}(\Delta) \mathbf{F}_P^*) = K \geq L \forall \Delta \neq 0$  in 1).
- 4) If  $\Delta$  is not an integer, we have also proved in 2) that  $\text{rank}(\mathbf{F}_2^T \mathbf{\Gamma}(\Delta) \mathbf{F}_P^*) = P \geq L$  if  $(N-P) \geq P$ . From 3) and 4), we conclude that the proposed design condition also satisfies (37) for JL's estimator.

## APPENDIX C

In this appendix, we will prove that the proposed pilot design conditions satisfy the consistency condition in the probabilistic sense, i.e., (47) satisfies (44) for the MLE, and (49) satisfies (48) for JL's estimator. Define  $\bar{\mathbf{R}}(\Delta) = \mathbf{U}_Z^H \mathbf{\Gamma}(\Delta) \mathbf{U}_P$ . First, we will prove it for the MLE.

- 1) For  $P \leq L$ ,  $\bar{\mathbf{R}}(\Delta)$  can be expressed as

$$\bar{\mathbf{R}}(\Delta) = \mathbf{A} \mathbf{T} \mathbf{D}_P \quad (98)$$

where  $\mathbf{A}$  is a full-rank  $(N-P) \times (N-P)$  matrix (since  $\text{rank}(\mathbf{U}_Z) = (N-P)$ ),  $\mathbf{D}_P$  is a full-rank  $P \times P$  matrix (since  $\text{rank}(\mathbf{U}_P) = P$ ), and

$$\mathbf{T} = \begin{bmatrix} \mathbf{f}_{n_1}^T \\ \mathbf{f}_{n_2}^T \\ \vdots \\ \mathbf{f}_{n_{N-P}}^T \end{bmatrix} \mathbf{\Gamma}(\Delta) [\mathbf{f}_{t_1}^*, \mathbf{f}_{t_2}^*, \dots, \mathbf{f}_{t_P}^*]. \quad (99)$$

If  $\Delta$  is a nonzero integer, the design condition  $\{t_k^{(l)}\} \neq \{t_k\}$  gives  $\mathbf{T} \neq \mathbf{0}$ , and since  $\mathbf{A}$  and  $\mathbf{D}_P$  are both full-rank square matrices, we have  $\bar{\mathbf{R}}(\Delta) \neq \mathbf{0}$ . If  $\Delta$  is not an integer, from (97), we know that  $\mathbf{T} \neq \mathbf{0}$  and hence, we have  $\bar{\mathbf{R}}(\Delta) \neq \mathbf{0}$ .

- 2) For  $P > L$ , from (91), we have

$$\bar{\mathbf{R}}(\Delta) = \mathbf{A} \begin{bmatrix} \mathbf{T}_1 \mathbf{D} \\ \mathbf{T}_2 \mathbf{D} \end{bmatrix}. \quad (100)$$

For a nonzero integer  $\Delta$ , the condition in (47) yields  $\mathbf{T}_1 = [\mathbf{e}_{l_1}, \mathbf{e}_{l_2}, \dots, \mathbf{e}_{l_K}, \mathbf{0}_{(N-P) \times (P-K)}] \mathcal{I}_{P \times P}$ . Also  $\mathbf{A}$  and  $\mathbf{D}$  are full-rank matrices, and hence,  $\bar{\mathbf{R}}(\Delta) \neq \mathbf{0}$ . For a noninteger  $\Delta$ , when  $1 \leq P \leq N/2$ , we have proved in Appendix B that  $\mathbf{T}_1$  is a full-rank matrix if  $(N-P) \geq P$ . Then, it follows that  $\bar{\mathbf{R}}(\Delta) \neq \mathbf{0}$ . When  $N/2 < P < (N+L)/2$ , we have  $\text{rank}(\mathbf{D}) > \text{Dimension}(\text{Nullspace}(\mathbf{T}_1))$ , and hence,  $\mathbf{T}_1 \mathbf{D} \neq \mathbf{0}$ , which yields  $\bar{\mathbf{R}}(\Delta) \neq \mathbf{0}$  for any noninteger  $\Delta$ . From 1) and 2), we conclude that (47) is the sufficient condition for (44). Next, we will prove it for JL's estimator.

- 3) If  $\Delta$  is a nonzero integer, the condition in (49) yields  $\{n_k^{(l)}\} \cap \{t_k\} \neq \emptyset$ , which guarantees (48).
- 4) If  $\Delta$  is not an integer,  $\mathbf{F}_2^T \mathbf{\Gamma}^H(\Delta) \mathbf{F}_P^*$  always has a full rank (see Appendix B), and hence, (48) is guaranteed. Thus, 3) and 4) complete our proof for JL's estimator.

## REFERENCES

- [1] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High-Speed Physical Layer in the 5 GHz Band*, IEEE Standard 802.11a, 1999.
- [2] *Broadband Radio Access Networks (BRAN); HIPERLAN Type 2; Physical (PHY) Layer Technical Specification*, ETSI TS 101 475 v1.1.1 (2000-04), 2000.
- [3] *IEEE Standards for Local and Metropolitan Area Networks, Part 16: Air Interface for Fixed Broadband Wireless Access Systems*, IEEE Standard 802.16a, 2004.
- [4] P. H. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Trans. Commun.*, vol. 42, no. 10, pp. 2908-2914, Oct. 1994.

- [5] T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Trans. Commun.*, vol. 45, no. 12, pp. 1614–1621, Dec. 1997.
- [6] M. Morelli and U. Mengali, "An improved frequency offset estimator for OFDM applications," *IEEE Commun. Lett.*, vol. 3, no. 3, pp. 75–77, Mar. 1999.
- [7] H. Minn, P. Tarasak, and V. K. Bhargava, "Some issues of complexity and training symbol design for OFDM frequency offset estimation methods based on BLUE principle," in *Proc. IEEE Veh. Technol. Conf.*, Apr. 2003, pp. 1288–1292.
- [8] H. Minn, V. K. Bhargava, and K. B. Letaief, "A robust timing and frequency synchronization for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 822–839, Jul. 2003.
- [9] H. Minn and P. Tarasak, "Improved maximum likelihood frequency offset estimation based on likelihood metric design," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2076–2086, Jun. 2006.
- [10] X. Ma, C. Tepedelenlioglu, G. B. Giannakis, and S. Barbarossa, "Non-data-aided carrier offset estimators for OFDM with null subcarriers: Identifiability, algorithm, and performance," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 12, pp. 2504–2515, Dec. 2001.
- [11] M. Ghogho, A. Swami, and G. B. Giannakis, "Optimized null-subcarrier selection for CFO estimation in OFDM over frequency-selective fading channels," in *Proc. IEEE Globecom*, 2001, vol. 1, pp. 202–206.
- [12] U. Tureli, P. J. Honan, and H. Liu, "Low-complexity nonlinear least squares carrier offset estimation for OFDM: Identifiability, diversity and performance," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2441–2452, Sep. 2004.
- [13] J. Lei and T.-S. Ng, "A consistent OFDM carrier frequency offset estimator based on distinctively spaced pilot tones," *IEEE Trans. Wireless Commun.*, vol. 3, no. 2, pp. 588–599, Mar. 2004.
- [14] 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.
- [15] 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.
- [16] H. Minn, X. Fu, and V. K. Bhargava, "Optimal periodic training signal for frequency offset estimation in frequency-selective fading channels," *IEEE Trans. Commun.*, vol. 54, no. 6, pp. 1081–1096, Jun. 2006.
- [17] S. Lin and D. J. Costello, Jr., *Error Control Coding*, 2nd ed. Englewood Cliffs, NJ: Pearson/Prentice-Hall, 2004.
- [18] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications Handbook*. New York: McGraw-Hill, 2002.
- [19] E. J. Weldon, Jr., "Difference-set cyclic codes," *Bell Syst. Tech. J.*, vol. XLV, no. 7, pp. 1045–1055, Sep. 1966.
- [20] D. Gordon, "La Jolla Cyclic Difference Set Repository," [Online]. Available: <http://www.ccrwest.org/diffsets.html>
- [21] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.



**Yinghui Li** (S'05) received the B.E. and M.S. degrees in electrical engineering from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2000 and 2003, respectively. She is currently working toward the Ph.D. degree in electrical engineering at the University of Texas at Dallas.

Her research interests are in the applications of statistical signal processing and training signal designs in synchronization, and channel estimation and detection problems in broadband wireless communications.



**Hlaing Minn** (S'99–M'01) received the B.E. degree in electronics from Yangon Institute of Technology, Yangon, Myanmar, in 1995, the M.Eng. degree in telecommunications from the Asian Institute of Technology (AIT), Pathumthani, Thailand, in 1997, and the Ph.D. degree in electrical engineering from the University of Victoria, Victoria, BC, Canada, in 2001.

He was with the Telecommunications Program in AIT as a Laboratory Supervisor during 1998. He was a Research Assistant from 1999 to 2001, and a Postdoctoral Research Fellow during 2002 in the Department of Electrical and Computer Engineering, University of Victoria. Since September 2002, he has been with the Erik Jonsson School of Engineering and Computer Science, University of Texas at Dallas, Richardson, as an Assistant Professor. His research interests include wireless communications, statistical signal processing, error control, detection, estimation, synchronization, signal design, and cross-layer design. He is an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS.



**Naofal Al-Dhahir** (S'89–M'90–SM'98) received the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, in 1990 and 1994, respectively, both in electrical engineering.

He was an Instructor at Stanford University in 1993. From August 1994 to July 1999, he was a Member of the Technical Staff at GE R&D Center, New York, where he worked on various aspects of satellite communication systems design and antijam GPS receivers. From August 1999 to July 2003, he was a Principal Member of Technical Staff at AT&T

Shannon Laboratory, Florham Park, NJ, where he worked on space-time coding and signal processing. Currently, he is an Associate Professor at the University of Texas at Dallas, Richardson. His current research interests include space-time coding and signal processing, OFDM, wireless networks, and digital subscriber line technology. He has authored over 160 journal and conference papers and holds 22 U.S. patents.

Dr. Al-Dhahir is a member of the IEEE SP4COM and SPTM technical committees. He served as Editor for IEEE TRANSACTIONS ON SIGNAL PROCESSING and IEEE COMMUNICATIONS LETTERS, and is currently an Editor for IEEE TRANSACTIONS ON COMMUNICATIONS. He served as Co-Chair of the Communication Theory Symposium at Globecom'04. He is a co-author of the book *Doppler Applications for LEO Satellite Systems* (New York: Springer, 2002). He is a co-recipient of the IEEE VTC Fall'05 Best Paper Award, the 2005 IEEE Signal Processing Society Young Author Best Paper Award, and the 2006 IEEE Donald G. Fink Best Paper Award.



**A. Robert Calderbank** (M'89–SM'97–F'98) received the B.Sc. degree in 1975 from Warwick University, Warwick, U.K., the M.Sc. degree in 1976 from Oxford University, Oxford, U.K., and the Ph.D. degree in 1980 from the California Institute of Technology, Pasadena, all in mathematics.

He is currently a Professor of Electrical Engineering and Mathematics at Princeton University, Princeton, NJ, where he directs the Program in Applied and Computational Mathematics. He joined Bell Telephone Laboratories as a Member

of Technical Staff in 1980, and retired from AT&T in 2003 as Vice President of Research. His research interests range from algebraic coding theory and quantum computing to the design of wireless and radar systems.

Dr. Calderbank served as Editor-in-Chief of the IEEE TRANSACTIONS ON INFORMATION THEORY from 1995 to 1998, and as Associate Editor for Coding Techniques from 1986 to 1989. He was a member of the Board of Governors of the IEEE Information Theory Society from 1991 to 1996, and began a second term in 2006. He was honored with the IEEE Information Theory Prize Paper Award in 1995 for his work on the Z4 linearity of Kerdock and Preparata Codes (joint with A. R. Hammons, Jr., P. V. Kumar, N. J. A. Sloane, and P. Sole), and again in 1999 for the invention of space-time codes (joint with V. Tarokh and N. Seshadri). He received the 2006 IEEE Donald G. Fink Prize Paper Award and the IEEE Millennium Medal, and was elected to the National Academy of Engineering in 2005.