[5] A. Mahajan and D. Teneketzis, "On the design of globally optimal communication strategies for real-time noisy communications systems with noisy feedback," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 4, pp. 580–595, May 2008.
[6] A. Ghasemi and E. S. Sousa, "Spectrum sensing in cognitive radio networks: Requirements, challenges and design trade-offs," *IEEE Commun. Mag*, vol. 46, no. 4, pp. 32–39, Apr. 2008.
[7] S. M. Mishra, A. Sahai, and R. W. Brodersen, "Cooperative sensing among cognitive radios," in *Proc. IEEE ICC*, Istanbul, Turkey, May 2006, pp. 1658–1663.
[8] G. Ganesan and Y. Li, "Cooperative spectrum sensing in cognitive radio, Part I: Two user networks," *IEEE Trans. Wireless Commun*, vol. 6, no. 6, pp. 2204–2213, Jun. 2007.
[9] G. Ganesan and Y. Li, "Cooperative spectrum sensing in cognitive radio, Part II: Multi user networks," *IEEE Trans. Wireless Commun*, vol. 6, no. 6, pp. 2204–2213, Jun. 2007.
[10] K. B. Letaief and W. Zhang, "Cooperative communications for cognitive radio networks," *Proc. IEEE*, vol. 97, no. 5, pp. 878–893, May 2009.
[11] D. Duan, L. Yang, and J. C. Principe, "Cooperative diversity of spectrum sensing for cognitive radio systems," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3218–3227, Jun. 2010.
[12] J. Ma, G. Zhao, and Y. Li, "Soft combination and detection for cooperative spectrum sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4502–4507, Nov. 2008.
[13] S. Zarrin and T. J. Lim, "Cooperative spectrum sensing in cognitive radios with incomplete likelihood functions," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3272–3281, Jun. 2010.
[14] S. Chaudhari, J. Lunden, V. Koivunen, and H. V. Poor, "Cooperative sensing with imperfect reporting channels: Hard decisions or soft decisions?" *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 18–28, Jan. 2012.
[15] A. L. Garcia, *Probability and Random Processes for Electrical Engineering*. Reading, MA: Addison-Wesley, 1994.
[16] J. G. Christiano and J. E. Hall, "On the n-th derivative of a determinant of the j-th order," *Math. Mag.*, vol. 37, no. 4, pp. 215–217, Sep. 1964.
[17] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1994.
[18] F. F. Digham, M.-S. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channels," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 21–24, Jan. 2007.
[19] D. Duan and L. Yang, "Cooperative sensing with ternary local decisions," in *Proc. IEEE ICASSP*, Kyoto, Japan, Mar. 2012, pp. 3677–3680.
[20] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[21] R. J. Schilling and S. L. Harris, *Applied Numerical Methods for Engineers Using MATLAB and C*. Pacific Grove, CA: Brooks/Cole, 2000.

Sequence Designs for Robust Consistent Frequency-Offset Estimation in OFDM Systems

Kilbom Lee, *Student Member, IEEE*,
Sung-Hyun Moon, *Member, IEEE*,
Seehyun Kim, and Inkyu Lee, *Senior Member, IEEE*

***Abstract*—In this paper, we derive the average pairwise error probability (PEP) of an integer carrier frequency-offset (CFO) estimator with consistent pilots in orthogonal frequency-division multiplexing (OFDM) systems and address several issues based on PEP analysis. In particular, the relationship between the PEP and consistent pilots is established in terms of a diversity gain and a shift gain. Based on the observations, we present new criteria for sequence designs. Simulation results show that the sequence developed from these criteria yields much reduced outliers compared with conventional sequences for consistent CFO estimation in frequency-selective fading channels.**

***Index Terms*—Carrier frequency offset (CFO), orthogonal frequency-division multiplexing (OFDM), pairwise error probability (PEP), synchronization.**

I. Introduction

Orthogonal frequency-division multiplexing (OFDM) is a popular modulation scheme in fading channels due to the effectiveness in dealing with fading-channel impairments such as cochannel interference and impulsive noise [1]. However, it is well known that the OFDM is quite sensitive to carrier frequency offset (CFO). The CFO can be divided into two parts. Fractional CFO (fCFO) refers to the CFO part whose value is less than one subcarrier spacing, whereas integer CFO (iCFO) represents the CFO with an integer multiple of the subcarrier spacing. The former destroys orthogonality among subcarriers and gives rise to a reduction in signal-to-noise ratio (SNR), whereas the latter causes a cyclic shift of subcarriers and results in detection errors. To guarantee reliable communication, the CFO should be correctly estimated so that the difference between the true CFO and the estimated CFO, which is defined as residual CFO (rCFO), is not greater than a certain level. The case where the rCFO is greater than the certain threshold is called an outlier; in this paper, we consider the outlier as the case with the rCFO greater than 0.5.

To reduce the outliers, several estimation schemes or pilot designs have been proposed [2]–[6]. Lei and Ng [4] addressed the relationship between the CFO and the periodogram of distinctively spaced pilot tones and proposed a maximum-likelihood estimator (MLE) for consistent CFO estimation over a frequency-selective fading channel. In [5], general pilot designs for the MLE were developed based

Manuscript received May 25, 2012; revised August 10, 2012; October 5, 2012; and October 24, 2012; accepted October 24, 2012. Date of publication November 20, 2012; date of current version March 13, 2013. This work was supported by the National Research Foundation of Korea funded by the Korean Ministry of Education, Science, and Technology under Grant 2010-0017909. The work was presented in part at the IEEE Global Communications Conference, Miami, FL, December 6–10, 2010. The review of this paper was coordinated by Dr. H. Lin.

K. Lee and I. Lee are with the School of Electrical Engineering, Korea University, Seoul 136-701, Korea (e-mail: bachhi@korea.ac.kr; inkyu@korea.ac.kr).

S.-H. Moon is with the Electronics and Telecommunications Research Institute, Daejeon 305-700, Korea (e-mail: sh.moon@etri.re.kr).

S. Kim is with the Department of Information and Telecommunications Engineering, University of Suwon, Hwaseong 445-743, Korea (e-mail: seehyun@suwon.ac.kr).

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

Digital Object Identifier 10.1109/TVT.2012.2228285

on the conventional sequences. These papers mainly focused on the conditions for pilots to ensure that the MLE has a unique solution for a noiseless channel. However, the conventional criteria for pilot designs are not generally robust to outliers in practical noisy channels, since the effect of the noise is neglected.

In contrast, Huang and Letaief [6] derived the pairwise error probability (PEP) of the MLE, which is defined as the probability that the estimated iCFO is not equal to the true iCFO, and showed that the consistent sequence[1] should have a large minimum Hamming distance to minimize the outliers. Moreover, the almost-perfect autocorrelation sequence was addressed, which maximizes the minimum Hamming distance [6]. Nevertheless, the sequence designed in [6] produces more outliers in frequency-selective channels, because the criteria are based on the PEP conditioned on the given channel realization.

In this paper, we first derive the average PEP of the MLE, which is averaged with respect to channel realizations. Unlike the conventional theorems in [5] and [6], we show that too large minimum Hamming distance leads to more outliers if the channel has insufficient frequency selectivity. Moreover, it is proven that the slope of the PEP at a high SNR is determined by the minimum value of the channel length and the minimum Hamming distance. As a result, we establish the relationship between the MLE performance and the consistent sequence in terms of a diversity gain and a shift gain. Based on these observations, we present new criteria for pilot designs to be suitable for the case where the received subcarriers are correlated. Simulation results show that consistent sequences satisfying the modified criteria can greatly reduce outliers compared with conventional sequences for consistent CFO estimation in frequency-selective fading channels.

Throughout this paper, normal letters represent scalar quantities, boldface lowercase letters indicate vectors, and uppercase letters designate matrices. We use $(\cdot)^T$, $(\cdot)^\dagger$, $\mathbb{E}(\cdot)$, $\|\cdot\|^2$, and $\langle\cdot\rangle_N$ for transpose, complex conjugate transpose, the expectation, the 2-norm operation, and the modulo N operation, respectively. In addition, $\mathbf{I}_n$ denotes an $n \times n$ identity matrix, and $\mathrm{diag}\{\mathbf{a}\}$ stands for a diagonal matrix whose diagonal elements are defined by $\mathbf{a}$.

In addition, $\mathcal{A} \setminus \mathcal{B}$ indicates the set $\{x \in \mathcal{A} | x \notin \mathcal{B}\}$.

II. Signal Model

In this paper, we consider OFDM systems with N subcarriers. We assume that all synchronization operations are perfectly made except for the CFO, which is denoted by Δf. We define the normalized CFO ϵ as $\epsilon \triangleq \Delta f / 1/T_s = \Delta f T_s$, where T_s stands for the OFDM symbol period. The normalized CFO ϵ can be divided into $\epsilon = \epsilon_i + \epsilon_f$, where ϵ_i and ϵ_f indicate an integer part and a fractional part ($|\epsilon_f| \le 1/2$), respectively.

Let us define the L-tap channel impulse response vector $\mathbf{h}$ as $\mathbf{h} = [h_0 h_1 \cdots h_{L-1}]^T$, where h_k has an independent and non-identically distributed (i.n.d.) complex Gaussian distribution with zero mean and variance $\sigma^2_{h_k}$, L represents the channel length, and we have $\sum_{k=0}^{L-1} \sigma^2_{h_k} = 1$. In addition, $\mathbf{H} = \sqrt{N}\mathrm{diag}\{\mathbf{F}_L \mathbf{h}\} = \mathrm{diag}\{H_0, H_1, \ldots, H_{N-1}\}$ denotes a diagonal matrix, where $\mathbf{F}_L$ indicates $\mathbf{F}_L = [\mathbf{f}_0 \mathbf{f}_1 \cdots \mathbf{f}_{L-1}]$ with $\mathbf{f}_i = 1/\sqrt{N}[1 e^{-j2\pi i/N} \cdots e^{-j2\pi i(N-1)/N}]^T$, and diagonal entry H_k for $k = 0, 1, \ldots, N-1$ is the channel gain at the kth subcarrier with $\mathbf{E}[|H_k|^2] = 1$. In addition, we define $\mathcal{C} = \{c_0, c_1, \ldots, c_{N_c - 1}\}$ as the set of indices of pilots where N_c is equal to the number of pilots.

Then, the received signal vector $\mathbf{r} = [r_0 r_1 \cdots r_{N-1}]^T$ after the cyclic prefix (CP) removal is given by

$$\mathbf{r} = \mathbf{\Gamma}(\epsilon)\mathbf{F}^\dagger \mathbf{H}\mathbf{x} + \mathbf{w}$$

[1]This is also called as the consistent pilots or distinctively spaced pilot tones [4].

where $\mathbf{\Gamma}(\epsilon) = \mathrm{diag}\{1, e^{j2\pi\epsilon/N}, \ldots, e^{j2\pi(N-1)\epsilon/N}\}$ represents a diagonal matrix whose diagonal entry accounts for the phase shift on the corresponding received signal sample, $\mathbf{F} = [\mathbf{f}_0 \mathbf{f}_1 \cdots \mathbf{f}_{N-1}]$ stands for the discrete Fourier transform (DFT) matrix, $\mathbf{x} = [X_0 X_1 \cdots X_{N-1}]^T$ is the transmitted signal vector with $|X_k|^2 = \varepsilon_x$ for $k \in \mathcal{C}$ and $X_k = 0$ otherwise, and $\mathbf{w} = [w_0 w_1 \cdots w_{N-1}]^T$ is equal to the complex additive white Gaussian noise vector with zero mean and covariance matrix $\sigma^2_w \mathbf{I}_N$.

Let us define $\mathcal{C}^{(\tau)}$ as the cyclic shifted set of $\mathcal{C}$ for $\tau \in \{0, 1, \ldots, N-1\}$, where the kth element of $\mathcal{C}^{(\tau)}$ is obtained as $c_k^{(\tau)} = \langle c_k + \tau\rangle_N$. Thereby, we denote $\mathbf{c}$ as the location column vector of length N whose kth element is $[\mathbf{c}]_k = 1$ for $k \in \mathcal{C}$ and 0 otherwise. The cyclic shifted vectors $\mathbf{c}^{(\tau)}$ in $\mathcal{C}^{(\tau)}$ are similarly defined. Then, the minimum Hamming distance of $\mathbf{c}$ is written as [5]

$$d_{\mathrm{H}} = 2\left(N_c - \max_{\tau \ne 0}\left(\mathbf{c}^T \mathbf{c}^{(\tau)}\right)\right) = 2(N_c - N_o) = 2N_t$$

where $N_o = \max_{\tau\ne 0}(\mathbf{c}^T\mathbf{c}^{(\tau)})$, and $N_t = N_c - N_o$. In this case, the size of the set $\mathcal{C} \setminus \mathcal{C}^{(\tau^*)}$ for $\tau^* = \arg\max_{\tau \ne 0}(\mathbf{c}^T \mathbf{c}^{(\tau)})$ becomes $N_t = d_{\mathrm{H}}/2$. For example, with $\mathbf{c} = [1\ 1\ 0\ 1]$, we have $N_c = 3$, $\tau^* = 1$, $N_o = 2$, $N_t = 1$, $d_{\mathrm{H}} = 2$, and $\mathcal{C} = \{1, 2, 4\}$.

III. Derivations of Pairwise Error Probability and Pilot Design Criteria

Here, we review the conventional PEP and criteria for pilot designs in [6]. Different from the conditional PEP in [6], we derive the PEP averaged over channel fadings and develop the feasible criteria.

A. Overview of Conventional PEP and the Design Criteria

We first assume that the true fCFO is obtained and removed from the received signal completely. Then, the conditional PEP that the erroneous iCFO $\bar{\epsilon}_i$ is chosen in MLE instead of ϵ_i is expressed by [6]

$$P(\epsilon_i \to \bar{\epsilon}_i | \mathbf{H}) \le \exp\left(-\frac{\eta(\bar{\epsilon}_i, \epsilon_i)}{4\sigma^2}\right). \tag{1}$$

Here, $\eta(\bar{\epsilon}_i, \epsilon_i)$ is defined as

$$\sum_{i+\Delta\epsilon \in \mathcal{C}, i \in \Gamma_y} |H_{i+\Delta\epsilon} X_{i+\Delta\epsilon}|^2 \tag{2}$$

where $\Delta\epsilon = \bar{\epsilon}_i - \epsilon_i$, and $\Gamma_y = \{0, 1, \ldots, N-1\} \setminus \mathcal{C}$.

Note that the PEP in (1) is conditioned on the given channel $\mathbf{h}$. To minimize the conditional PEP, we need to maximize the minimum of $\eta(\bar{\epsilon}_i, \epsilon_i)$ for $\bar{\epsilon}_i \ne \epsilon_i$. Thus, the number of elements in (2) should be as large as possible. Then, the criteria for robust pilot designs are summarized as follows [6]:

1) The sequence should have a good distance distribution in the sense that the cyclic-shift versions with small distance to the original binary sequence are as few as possible.
2) Zeros in the binary sequence should be as far apart as possible to ensure frequency-domain diversity.

B. Derivation of Average PEP

With $\epsilon = \epsilon_i$, the MLE is written by [4]

$$\tilde{\epsilon}_i = \arg\max_{\hat{\epsilon}_i} \left\|\mathbf{F}^T_{N_c}\mathbf{\Gamma}^\dagger(\hat{\epsilon}_i)\mathbf{r}\right\|^2 = \arg\max_{\hat{\epsilon}_i} \sum_{k\in\mathcal{C}} \left|R_{\langle k+\hat{\epsilon}_i\rangle_N}\right|^2 \tag{3}$$

where $\tilde{\epsilon}_i$ indicates the estimated iCFO value, and $\mathbf{F}_{N_c}$ and R_k are defined by $\mathbf{F}_{N_c} = [\mathbf{f}_{c_0} \mathbf{f}_{c_1} \cdots \mathbf{f}_{c_{N_c-1}}]$ and $R_k = 1/\sqrt{N} \times$

$\sum_{n=0}^{N-1} r_n e^{-j2\pi kn/N}$, respectively. Here, we assume that $N_t \geq 1$, and then, $\mathbf{x}$ becomes a consistent[2] sequence. Otherwise, the MLE in (3) always has several maxima even for a noiseless channel, which indicates that the estimator is not consistent.

Since the channel and the noise are zero mean complex jointly Gaussian and the training sequence is deterministic, the received signal has zero mean complex jointly Gaussian components. Thus, the distribution of $|R_k|^2$ is exponential, whose mean is

$$\mathbb{E}\left[\left|R_{\langle k+\epsilon_i\rangle_N}\right|^2\right] = \begin{cases} \varepsilon_x + \sigma_w^2, & \text{for } k \in \mathcal{C} \\ \sigma_w^2, & \text{otherwise.} \end{cases}$$

Based on the iCFO estimator (3), the PEP can be represented by

$$P(\epsilon_i \to \bar{\epsilon}_i) = P\left(\sum_{k\in\mathcal{C}} \left|R_{\langle k+\epsilon_i\rangle_N}\right|^2 < \sum_{k\in\mathcal{C}} \left|R_{\langle k+\bar{\epsilon}_i\rangle_N}\right|^2\right). \quad (4)$$

Subtracting the 2-norm of the received pilots common in both sides, (4) can be transformed into

$$P(\epsilon_i \to \bar{\epsilon}_i) = P(Q < Z) \quad (5)$$

where $Q = \sum_{k\in\overline{\mathcal{C}}} |R_{\langle k+\epsilon_i\rangle_N}|^2$ and $Z = \sum_{k\in\overline{\mathcal{B}}} |R_{\langle k+\bar{\epsilon}_i\rangle_N}|^2$. Here, $\overline{\mathcal{C}}$ and $\overline{\mathcal{B}}$ are defined as $\overline{\mathcal{C}} = \{\bar{c}_0, \bar{c}_1, \ldots, \bar{c}_{N_t-1}\} = \mathcal{C} \setminus \mathcal{C}^{(\bar{\epsilon}_i - \epsilon_i)}$ and $\mathcal{C} \setminus \mathcal{C}^{(\epsilon_i - \bar{\epsilon}_i)}$, respectively. In addition, those sizes are assumed to be N_t to consider the worst case. Comparing (4) and (5) shows that $R_{\langle k+\bar{\epsilon}_i\rangle_N}$ in Z includes just noise. Since Z is a summation of the 2-norm of the independent identically distributed (i.i.d.) noise terms, Z has a chi-square distribution with N_t degrees of freedom (DOF), whose probability density function (pdf) is given by

$$f_Z(z) = \frac{1}{\sigma_w^{2N_t}(N_t-1)!} z^{N_t-1} \exp\left(-\frac{z}{\sigma_w^2}\right).$$

In contrast, the received subcarriers in Q are correlated by the channel, and the pdf of Q is not a chi-square distribution. To derive the pdf of Q, we first define the received vector as $\dot{\mathbf{r}} = [R_{\langle \bar{c}_0+\epsilon_i\rangle_N} \cdots R_{\langle \bar{c}_{N_t-1}+\epsilon_i\rangle_N}]^T$ and denote the autocorrelation matrix of $\dot{\mathbf{r}}$ as $\mathbf{A}_R = \mathbb{E}[\dot{\mathbf{r}}\dot{\mathbf{r}}^\dagger]$, which can be eigendecomposed into

$$\mathbf{A}_R = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^\dagger \quad (6)$$

where $\mathbf{U}$ indicates a unitary matrix, an $N_t \times N_t$ matrix $\boldsymbol{\Lambda}$ is equal to $\mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_d, 0, 0, \ldots, 0\}$ with λ_k being the kth eigenvalue of $\mathbf{A}_R$, and d stands for the rank of $\mathbf{A}_R$ with $d \leq N_t$.

Then, Q is distributed as $\sum_{k=1}^{d} G_k$, where G_k represents the exponential random variable with $f_{G_k}(g) = 1/\lambda_k \exp(-g/\lambda_k)$ [7]. Here, we see that the distribution of Q can be expressed by the weighted sum of d independent exponential variables. Now, we denote the Taylor series of $f_{G_k}(g)$ as $\lambda_k^{-1} + o(g)$. Then, by using [8, Th. 2] and $\lambda_k^{-1} + o(g)$, the series expression of the pdf of Q around zero is written as

$$f_Q(q) = \frac{\prod_{k=1}^{d} \lambda_k^{-1}}{\Gamma(d)} q^{d-1} + o(q^d) \quad (7)$$

where $\Gamma(\cdot)$ indicates the Gamma function. Note that the series expression of the pdf around zero is generally used to analyze the performance at a high SNR, because the diversity and shift gains for average PEP only depend on the behavior of the pdf around the origin (i.e., $q \to 0$) [9].

[2]In this paper, the consistency accounts for *the consistency in the probabilistic sense* as in [5].

For simple calculations, we evaluate $P(Q > Z) = 1 - P(Q < Z)$ as

$$\begin{aligned} P(Z < Q) &= \int_0^\infty P(Z < q) f_Q(q)\, dq \\ &= \int_0^\infty \left[1 - \sum_{n=0}^{N_t-1} e^{-q/\sigma_w^2} \left(\frac{q}{\sigma_w^2}\right)^n \frac{1}{n!}\right] f_Q(q)\, dq \\ &\approx 1 - \sigma_w^{2d} \prod_{k=1}^{d} \lambda_k^{-1} \sum_{n=0}^{N_t-1} \frac{\Gamma(n+d)}{\Gamma(d)\Gamma(n+1)} \end{aligned} \quad (8)$$

where the approximation in (8) results from ignoring $o(q^d)$ in (7), because the first term $o(q^{d-1})$ is dominant as $q \to 0$. Taking the log operation, then the PEP is evaluated as

$$\log P(\epsilon_i \to \bar{\epsilon}_i) \approx -d \log\left(\frac{1}{\sigma_w^2}\right) - \sum_{k=1}^{d} \log \lambda_k + \log \sum_{n=0}^{N_t-1} \binom{n+d-1}{n}. \quad (9)$$

Here, we can easily verify that d accounts for a diversity gain in the PEP curve, and d, N_t, and λ_k determine the horizontal shift in the curve.

C. Observations in the Average PEP

Based on the average PEP, we provide three lemmas and corresponding observations to clarify the relationship between the consistent sequence and the MLE performance.

Lemma 1: The diversity gain of the MLE for L-tap fading channels is given by

$$\lim_{\mathrm{SNR}\to\infty} -\frac{\log P(\epsilon_i \to \bar{\epsilon}_i)}{\log \mathrm{SNR}} = \min(L, N_t).$$

Proof: From the definition of the diversity gain [9], [10], we apply the high-SNR approximation to $\mathbf{A}_R$ in (6). In this case, the rank d of $\mathbf{A}_R$ is obtained as $\min(L, N_t)$. Therefore, the diversity gain of the MLE for L-tap fading channels is equal to $\min(L, N_t)$. ■

From the given lemma, we see that the consistent sequence with $N_t < L$ results in a diversity loss in the MLE, whereas a diversity gain is not greater than L even if $N_t > L$. Meanwhile, the average PEP (9) shows that the third term $\log \sum_{n=0}^{N_t-1} \binom{n+d-1}{n}$ increases with N_t. Hence, if N_t is greater than L, the PEP curve is shifted to the right, whereas the diversity gain is fixed to L. This is due to the fact that the DOF of Z is equal to N_t, whereas that of Q is given by $\min(N_t, L)$. As a result, N_t determines the shift gain and the diversity gain.

In the following two lemmas, we show that the shift gain $\sum_{k=1}^{d} \log \lambda_k$ in (9) depends on the ratio of N_o to N_c and the distribution of pilots in $\overline{\mathcal{C}}$.

Lemma 2: Defining $P_T = N_c \varepsilon_x$ as the total transmit power, the expectation of Q is bounded by

$$P_T\left(1 - \frac{N_o}{N_c}\right) \leq \mathbb{E}(Q) < P_T.$$

Proof: From (6), $\mathbb{E}[Q] = \mathbb{E}[\dot{\mathbf{r}}^\dagger \dot{\mathbf{r}}] = \mathrm{Tr}(\mathbf{A}_R) = \sum_{k=1}^{d} \lambda_k = N_t(\varepsilon_x + \sigma_w^2)$. Then, the ratio of $\mathbb{E}(Q)$ to P_T can be expressed as

$$\frac{\mathbb{E}(Q)}{P_T} = \frac{N_t\left(\varepsilon_x + \sigma_w^2\right)}{N_c \varepsilon_x} = \frac{N_t}{N_c} + \frac{N_t \sigma_w^2}{N_c \varepsilon_x} \geq \frac{N_t}{N_c} = 1 - \frac{N_o}{N_c}.$$

■

Here, it can be checked that a lower bound of $\mathbb{E}(Q) = \sum_{k=1}^{d} \lambda_k$ decreases as N_o/N_c increases. Therefore, we expect that the growth of N_o reduces $\sum_{k=1}^{d} \log \lambda_k$ in (9), which makes the PEP curve shift to the right. Meanwhile, most of the conventional consistent sequences in [5] and [6] such as the almost-perfect autocorrelation sequence have large N_o and N_t. In this case, performance loss due to the reduced $\sum_{k=1}^{d} \log \lambda_k$ is inevitable.

Lemma 3: For a given N_c and $N_t \geq L$, uniformly distributed pilots in $\overline{\mathcal{C}}$ maximize the shift gain $\sum_{k=1}^{d} \log \lambda_k$ in (9) for equal-power fading channels.

Proof: From the definition of a coding (shift) gain, we assume $\sigma_w^2 \to 0$ [9]. Moreover, ϵ_i is set to zero without loss of generality, because iCFO only causes a cyclic shift of subcarriers. Since $\sum_{k=1}^{d} \lambda_k = N_t \varepsilon_x = (N_c - N_o)\varepsilon_x$ is constant, $\prod_{k=1}^{d} \lambda_k$ is maximized when all λ_k are the same. Meanwhile, we define an $N \times N$ diagonal matrix $\tilde{\mathbf{\Phi}}$, where $[\tilde{\mathbf{\Phi}}]_{i,i} = 1$ for $i \in \overline{\mathcal{C}}$ and $[\tilde{\mathbf{\Phi}}]_{i,i} = 0$ otherwise. Extracting the rows whose indices are contained in $\overline{\mathcal{C}}$ from $\tilde{\mathbf{\Phi}}$ yields $\mathbf{\Phi}$ of size $N_t \times N$ with $\mathbf{\Phi}\mathbf{\Phi}^T = \mathbf{I}_{N_t}$.

With $N_t \geq L$, $\mathbf{A}_R$ can be expressed as

$$\begin{aligned}
\mathbf{A}_R &= \mathbf{E}[\dot{\mathbf{r}}\dot{\mathbf{r}}^\dagger] \\
&= N \mathrm{diag}\{\mathbf{\Phi x}\}\mathbf{\Phi F}_{N_t}\mathbf{E}[\tilde{\mathbf{h}}\tilde{\mathbf{h}}^\dagger]\mathbf{F}_{N_t}^\dagger \mathbf{\Phi}^T \mathrm{diag}\{\mathbf{x}^\dagger \mathbf{\Phi}^T\} \\
&= \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\dagger
\end{aligned}$$

where $\tilde{\mathbf{h}}$ is a zero-padded vector of size N_t as $\tilde{\mathbf{h}} = [h_0 h_1 \cdots h_{L-1}, 0, \cdots, 0]^T$, and $\mathbf{\Lambda}$ and $\mathbf{U}$ are given by $\mathrm{diag}\{N_t \varepsilon_x \sigma_{h_0}^2, \cdots, N_t \varepsilon_x \times \sigma_{h_{L-1}}^2, 0, \cdots, 0\}$ and $\sqrt{N/N_t\varepsilon_x}\mathrm{diag}\{\mathbf{\Phi x}\}\mathbf{\Phi F}_{N_t}$, respectively. Here, we see from [11] that $\mathbf{U}^\dagger\mathbf{U} = \mathbf{U}\mathbf{U}^\dagger = \mathbf{I}_{N_t}$ if and only if pilots in $\overline{\mathcal{C}}$ are equispaced and equipowered. In this case, $\mathbf{U}$ becomes a unitary matrix, and thus, the diagonal elements of $\mathbf{\Lambda}$ are the eigenvalues of $\mathbf{A}_R$. As a result, for equal-power fading channels (i.e., $\sigma_{h_i}^2 = \sigma_{h_j}^2$, $\forall i \neq j$), $\prod_{k=1}^{L} \lambda_k$ is maximized. ■

It is well known that the white sequence, which comprises equispaced and equipowered pilots, is optimal if the channel state information is not known to the transmitter, because the worst case asymptotic Cramer–Rao bound (CRB) is minimized [12]. However, it is impossible to make such a consistent sequence since both uniformity and consistency stand in contradiction to each other. Note that the white sequence is not consistent because there is nonzero τ, which satisfies $\mathbf{c}^T\mathbf{c}^{(\tau)} = N_c$. Nevertheless, almost uniformly distributed pilots provide a large $\prod_{k=1}^{d} \lambda_k$ (i.e., $\sum_{k=1}^{d} \log \lambda_k$), which will be confirmed in the simulation section. As a result, it can be easily seen how pilot designs affect the MLE performance in frequency-selective fading channels by analyzing the derived PEP with respect to a diversity gain and a shift gain.

D. Modified Criteria for Consistent CFO Estimation Against Outliers

Based on the observations made in the previous section, we now modify the conventional criteria to ensure robustness to outliers in frequency-selective fading channels. From Lemma 1, we learn that N_t should be set to a value larger than L to achieve the diversity gain of L. In contrast, too large N_t yields performance degradations when the channel has insufficient frequency selectivity (i.e., $L < N_t$). Moreover, Lemma 2 indicates that N_o/N_c should be as small as possible. Finally, from Lemma 3, the pilot tones should be located as uniform as possible to achieve a large $\prod_{k=1}^{L} \lambda_k$. In conclusion, our modified criteria are simply summarized as follows:

Criterion 1: N_t should be larger than L.

Criterion 2: N_t and N_o should be set as small as possible.

Criterion 3: A distribution of pilot tones should be as uniform as possible.

TABLE I
CONSISTENT SEQUENCES

Sequence	Pilot Index	N_t	N_o	$\sum_{k=1}^{d} \log_2 \lambda_k$
A-Seq	{0, 8, 17, 27, 38, 50}	5	1	13.55
B-Seq	{0, 2, 6, 14, 30, 53}	5	1	7.73
C-Seq	{0, 18, 37, 57}	3	1	10.79
EM-Seq	A4E2F28C20FD59BA	15	17	10.5
AP-Seq	0C6A01B2F3957E4D	16	15	10.87

Note that Criteria 1 determines the diversity gain, whereas Criterion 2 and 3 address the shift gain, as will be shown in the simulation section.

IV. SIMULATION RESULTS

Here, we present the simulation results of consistent CFO estimation to show the efficacy of our modified criteria. The number of subcarriers, CP samples, and the simulation runs are set to $N = 64$, $N_g = 16$, and 100 000 000, respectively. For simplicity, we refer to the sequence that meets all the criteria as A-Seq. Moreover, we call the sequences satisfying Criterion 1 and 3 and Criterion 2 and 3 as B-Seq and C-Seq, respectively. In addition, the extended m-sequence in [5] and the almost-perfect autocorrelation sequence in [6] are denoted by EM-Seq and AP-Seq, respectively. The indices of pilots of the given sequences are given by Table I. Here, the hexadecimal digits can be converted to binary digits, where "1" and "0" correspond to a pilot tone and a null subcarrier, respectively. A-Seq, B-Seq, and C-Seq are generated by computer search based on the modified criteria, and thus, the sequences are not unique. Meanwhile, to demonstrate the efficacy of the MLE with the proposed sequences, we employ a joint ML estimator for the channel impulse response and iCFO. In addition, its low-complexity version developed in [13] is considered. For brevity, these are referred to as JML and LC-JML, respectively. Note that the JML and the LC-JML can estimate the iCFO through uniformly distributed pilots by exploiting the knowledge of the pilot values, unlike the MLE.

In Figs. 1 and 2, we assume 4-tap equal-power fading channels, and the normalized CFO is randomly selected in $\{0, 1, \ldots, 63\}$. Fig. 1 shows the error probability of consistent iCFO estimation for various sequences. It is shown that A-Seq exhibits the smallest outliers, and its diversity gain is 4, which matches with our analysis. In contrast, B-Seq shows a much lower shift gain, although N_t and N_o of B-Seq are identical to those of A-Seq. Moreover, Table I indicates that B-Seq has the smallest value of $\sum_{k=1}^{d} \log_2 \lambda_k$, whereas A-Seq has the highest value. Therefore, we see that a shift gain of the MLE strongly depends on a distribution of pilots. On the other hand, the diversity gain of AP-Seq is almost identical to that of A-Seq, whereas the performance is 4 dB worse than A-Seq in most SNR regions due to large N_t and N_o. In addition, its performance is also worse than that of C-Seq in the SNR less than 15 dB, although the latter may yield an ambiguity problem.[3] This can be explained by the fact that

[3]In noiseless systems, [4] and [5] proved that a solution of the MLE is unique for any channel realization only when $N_t \geq L$ Otherwise, the MLE may have several maxima for a particular channel realization, which results in complete failure for all SNR values.

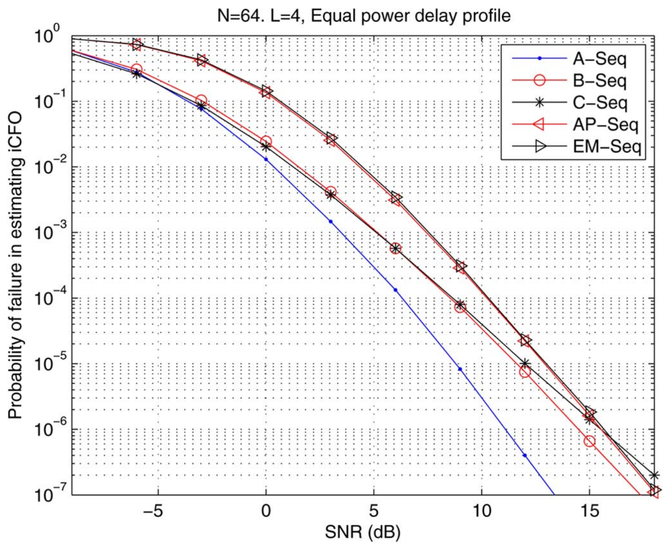

Fig. 1. Probability of failure in estimating iCFO.

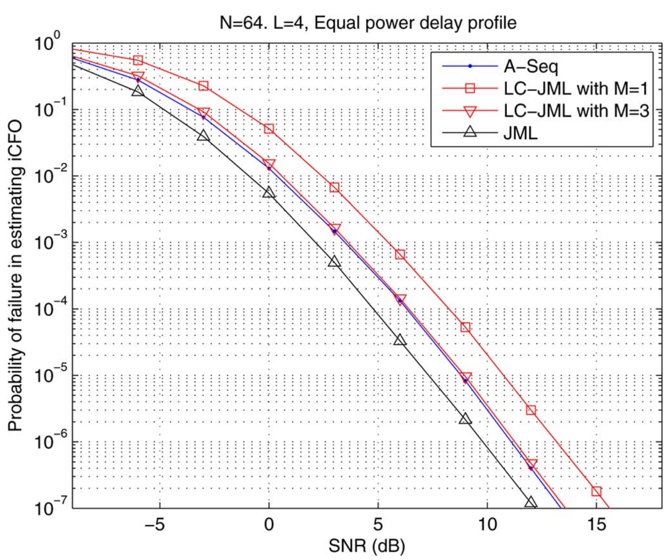

Fig. 2. Probability of failure in estimating iCFO.

TABLE II
COMPUTATIONAL COMPLEXITY

Algorithm	real products
JML	$N(4N(L-1)-2L)$
LC-JML	$2NM(2N-M)$
MLE	$2N$

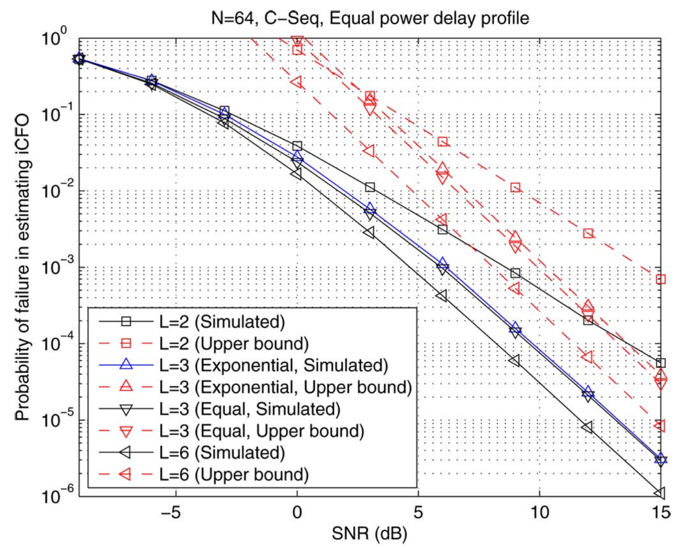

Fig. 3. Probability of failure in estimating iCFO.

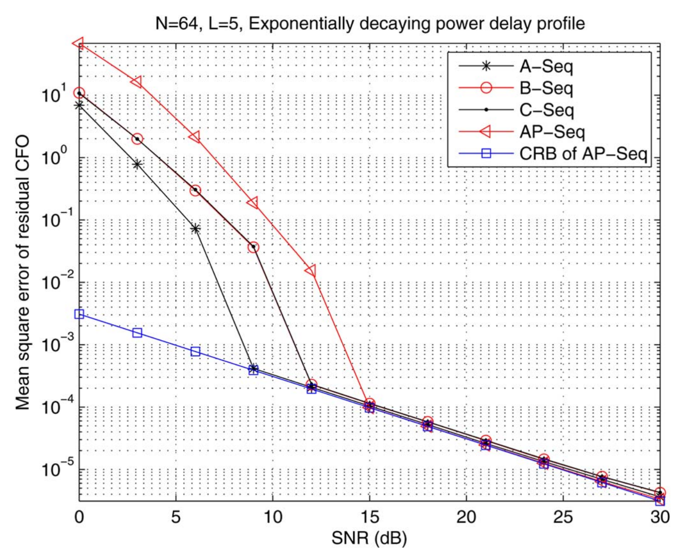

Fig. 4. Mean square error of rCFO.

the probability of encountering such channel realizations is very low, and thus, the effect is masked by the noise at a low and moderate SNR. Hence, it is shown that the larger minimum Hamming distance does not guarantee better performance. On the other hand, we check that B-Seq and AP-Seq outperform C-Seq at the high SNR, because C-Seq suffers from a diversity loss. It should be noted that B-Seq and AP-Seq can achieve the diversity gain 4, although those eigenvalues are different from each other. As a result, for a given L, the MLE performance of each consistent sequence can be well explained by our modified criteria.

Fig. 2 shows the comparison of the MLE with the A-Seq, the JML, and the LC-JML in terms of the failure probability for estimating iCFO, where M stands for the iteration numbers. For fair comparison, all algorithms use just one training sequence. In this plot, we see that the JML and the LC-JML exhibit the same diversity gain as the MLE with A-Seq, whereas the JML is 1.2 dB better than the MLE. In addition, it appears that both the MLE with A-Seq and the LC-JML with $M = 3$ have virtually the same accuracy, whereas the MLE is clearly superior than the LC-JML with $M = 1$. Meanwhile, Table II shows that the number of real products of the JML and the LC-JML is given by $O(N^2)$, whereas that of the MLE is equal to $2N$. For example of $N = 64$, the complexity of the MLE is only 0.1% and 0.3% compared with JML and LC-JML with $M = 3$, respectively.

As a result, the MLE with our proposed sequence may be preferable for practical systems over the JML, since the MLE requires much lower computational complexity at the expense of a small shift loss.

Fig. 3 shows the failure probability of C-Seq for different channel lengths, where equal-power fading channels are employed except for the graph labeled "Exponential." In addition, the upper bounds are computed by using the equation $(N-1)\sigma_w^{2d}\prod_{k=1}^{d}\lambda_k^{-1} \times \sum_{n=0}^{N_t-1}\binom{n+d-1}{n}$, where the values of λ_k are obtained through numerical simulations. The graphs with $L = 3$ confirm that the C-Seq achieves a diversity gain of $\min(L, N_t)$ regardless of the power delay profile, whereas equal-power fading channels provide a slightly better shift gain than the exponentially decaying channel. Moreover, it is verified that the C-Seq with $L = 6$ exhibits a higher shift gain than that with $L = 3$, but the diversity gain still remains as 3. As a result, if the channel length is not considered in pilot designs, the MLE cannot fully exploit the DOF of the channel. Note that these results can be also verified from the upper bounds.

To confirm our modified criteria for consistent CFO estimation, the results in Fig. 4 are simulated for 5-tap exponentially decaying fading channels with $\epsilon = 30.24$ and the simulation runs of 100 000. In addition, the CRB of the AP-Seq is obtained from [14]. We see that the MLE approaches the CRB at a high SNR, and various consistent sequences exhibit different performance. It should be noted that the large mean square error (MSE) in low- and moderate-SNR regions results from outliers. Similar to the previous simulation results, C-Seq shows reduced outliers than B-Seq and AP-Seq in low- and moderate-SNR regions, whereas A-Seq exhibits the smallest outliers. As a result, we confirm that A-Seq provides the best performance over other sequences for consistent CFO estimation. Hence, it is shown that our modified criteria are also effective for consistent CFO estimation.

V. Conclusion

In this paper, we have derived the average PEP of the MLE and established the relationship of the PEP and the consistent sequence in terms of a diversity gain and a shift gain. Then, we have presented the criteria suitable for the case where the subcarriers are correlated. Simulation results show that the consistent sequences based on our modified criteria produce much reduced outliers over conventional sequences for consistent CFO estimation in frequency-selective fading channels.

References

[1] J. A. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 25, no. 5, pp. 5–14, May 1990.

[2] K. Lee and I. Lee, "Robust pilot designs for consistent frequency offset estimation in OFDM systems," in *Proc. IEEE GLOBECOM*, Dec. 2010, pp. 1–5.

[3] K. Lee, S.-H. Moon, and I. Lee, "Low-complexity leakage-based carrier frequency offset estimation techniques for OFDMA uplink systems," in *Proc. IEEE GLOBECOM*, Dec. 2010, pp. 1–5.

[4] 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.

[5] Y. Li, H. Minn, N. Al-Dhahir, and A. R. Calderbank, "Pilot designs for consistent frequency-offset estimation in OFDM systems," *IEEE Trans. Commun.*, vol. 55, no. 5, pp. 864–877, May 2007.

[6] D. Huang and K. B. Letaief, "Carrier frequency offset estimation for OFDM systems using subcarriers," *IEEE Trans. Commun.*, vol. 54, no. 5, pp. 813–823, May 2006.

[7] Q. T. Zhang and D. P. Liu, "A simple capacity formula for correlated diversity Rician fading channels," *IEEE Commun. Lett.*, vol. 6, pp. 481–483, Nov. 2002.

[8] Z. Fang, X. Bao, L. Li, and Z. Wang, "Performance analysis and power allocation for amplify-and-forward cooperative networks over Nakagami-m fading channel," *IEICE Trans. Commun.*, vol. E92-B, no. 3, pp. 1004–1012, Mar. 2009.

[9] Z. Wang and G. B. Giannakis, "A simple and general parameterization quantifying performance in fading channels," *IEEE Trans. Commun.*, vol. 51, no. 8, pp. 1389–1398, Aug. 2003.

[10] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[11] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Trans. Signal Process.*, vol. 51, no. 6, pp. 1615–1624, Jun. 2003.

[12] 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.

[13] M. Morelli and M. Moretti, "Integer frequency offset recovery in OFDM transmissions over selective channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5220–5226, Dec. 2008.

[14] Y. Li, H. Minn, and J. Zeng, "An average Cramer–Rao bound for frequency offset estimation in frequency-selective fading channels," *IEEE Trans. Commun.*, vol. 9, no. 3, pp. 871–875, Mar. 2010.

Variational-Inference-Based Data Detection for OFDM Systems With Imperfect Channel Estimation

Feng Li, Zongben Xu, and Shihua Zhu

***Abstract*—This paper studies the problem of joint estimation of data and channels for orthogonal frequency-division multiplexing (OFDM) systems using variational inference. The proposed methods are used to combat imperfect channel estimation at the receiver since it can degrade system performance seriously. The proposed methods simplify the maximum *a posteriori* (MAP) scheme based on the theory of variational inference and formulate an optimization problem using variational free energy. The channel state information (CSI) and data are dealt with jointly and iteratively. The proposed schemes offer a variety of solutions for getting soft information when turbo equalization is implemented for coded systems. The effectiveness of the new approach is demonstrated by Monte Carlo simulations.**

***Index Terms*—Channel imperfections, orthogonal frequency-division multiplexing (OFDM), variational inference.**

I. Introduction

Orthogonal frequency-division multiplexing (OFDM) is being investigated as a promising solution of physical layers for current and future wireless communication systems. However, much attention should be paid to several problems that challenge the practical utilization of OFDM, including channel estimation, time-and-frequency offset, phase noise, peak-to-average power ratio, and data detection [1]. In this paper, we consider the problem of data detection with imperfect channel estimation.

Anderson [2] studies the relationship between channel fade and its estimate. The model used in [2] is specialized to three typical channel estimation approaches by Annavajjala *et al.* in [3]. There have been many effects in analyzing OFDM with imperfect channel estimation, e.g., in [4] and [5]. Song *et al.* [4] show us that, compared with other parameter imperfections, inherent channel estimation imperfection is the main reason that degrades the system performance. Krondorf and Fetweiss [5] investigate the Alamouti space–time-coded OFDM performance under receiver impairments, and it is demonstrated that channel uncertainty weights more heavily than carrier frequency offset and I/Q imbalance against system performance. Since imperfect

Manuscript received April 22, 2012; revised August 28, 2012; accepted November 30, 2012. Date of publication December 5, 2012; date of current version March 13, 2013. This work was supported in part by the National Natural Science Foundation of China under Grant 11131006 and Grant 61075054, by the International Cooperation and Exchanges Project under Grant S2010GR0902, and by the National Major Special Projects in Science and Technology of China under Grant 2011ZX03003-002-03. The review of this paper was coordinated by Prof. G. Bauch.

F. Li and S. Zhu are with the Department of Information and Communication Engineering, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: lf1981@mail.xjtu.edu.cn; szhu@mail.xjtu.edu.cn).

Z. Xu is with the Institute for Information and System Sciences, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: zbxu@mail.xjtu.edu.cn).

Digital Object Identifier 10.1109/TVT.2012.2231972