Blind Maximum Likelihood CFO Estimation for OFDM Systems via Polynomial Rooting

Feifei Gao and A. Nallanathan, *Senior Member, IEEE*

***Abstract*—The blind carrier frequency offset estimation problem has been well studied by exploiting the virtual carriers existing in practical orthogonal frequency division multiplexing transmissions. A highly efficient approach by rooting a polynomial has been proposed in the literature. However, this rooting method is suboptimal when noise is present. In this letter, we propose an improved polynomial rooting method that is shown to be the maximum likelihood estimator for both the noisy and the noise-free case.**

***Index Terms*—Blind carrier frequency offset (CFO) estimation, orthogonal frequency division multiplexing (OFDM), polynomial rooting, virtual carriers.**

I. Introduction

ORTHOGONAL frequency division multiplexing (OFDM) [1] is a promising candidate for next-generation wireless communication due to its high data rate, high spectral efficiency, and robustness to frequency-selective channels. However, the presence of carrier frequency offset (CFO) caused by the mismatch in oscillators or Doppler effect destroys the orthogonality among subcarriers and results in a severe degradation in bit-error rate (BER) performance [2]. Therefore, CFO must be compensated before channel estimation and coherent data detection.

A class of blind CFO estimation methods has been proposed in [3]–[6], where the existence of null subcarriers in practical OFDM systems is exploited. This method has two types of implementation. One is the MUSIC-like searching approach [3], [4] that is proved to be the maximum likelihood (ML) estimator for CFO estimation [5], [6]. The other is the search-free approach that was proposed in [3] where the root-MUSIC-like polynomial rooting is exploited. Since the root-MUSIC algorithm computes the roots directly from the cost function, the resulting solution is not the ML estimate in the presence of noise.

In this letter, we propose an improved polynomial rooting method. Instead of rooting the cost function directly, the basic idea behind is to root the first-order derivative of the cost function. This approach can yield an ML estimate of CFO, as will be shown later. One can see from the simulation results that the performance of the direct rooting method (DIRM) reaches that of the derivative rooting method (DERM) asymptotically at high signal-to-noise ratio (SNR), whereas, its performance loss is obvious at low SNR.

Manuscript received August 15, 2005; revised September 29, 2005. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Stefano Galli.

The authors are with the Department of Electrical and Computer Engineering, National University of Singapore, 119260 Singapore (e-mail: g0403702@nus.edu.sg; elena@nus.edu.sg).

Digital Object Identifier 10.1109/LSP.2005.861583

II. Problem Formulation

A. System Model

For a consistent study of the problem, we use the OFDM model provided in [3] and [4]. Here, we assume that the subcarriers with index 1 to P from total N subcarriers are used for data transmissions. Let

$$\mathbf{s}(k) = [s_1(k), s_2(k), \ldots, s_P(k)]^T, \qquad k = 1, \ldots, K \quad (1)$$

denote the kth block of data to be transmitted. The time-domain signals can be obtained via N points IDFT operation as

$$\mathbf{x}(k) = [x_1(k), x_2(k), \ldots, x_N(k)]^T = \mathbf{W}_p\mathbf{s}(k) \quad (2)$$

where

$$\mathbf{W}_P = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_P] \quad (3)$$

contains the first P columns of the IDFT matrix whose (u,v)th entry is given by $(1/\sqrt{N})e^{j2\pi(u-1)(v-1)/N}$. Clearly, both $\mathbf{W}_P$ and its orthogonal complement matrix, defined as

$$\mathbf{W}_\perp = [\mathbf{w}_{P+1}, \ldots, \mathbf{w}_N] \quad (4)$$

are known *a priori*. At the transmitter, the symbols are sent at a rate of $1/T_s$, and a cyclic prefix (CP) of length G is added in front of each OFDM block. Let $\mathbf{h} = [h_0, h_1, \ldots, h_L]$ denote the equivalent discrete channel impulse response and Δf represent the carrier frequency offset. Define new notations

$$\phi_0 = 2\pi\Delta f T_s \quad (5)$$

$$\mathbf{\Phi}_0 = \mathrm{diag}\left\{1, e^{j\phi_0}, \ldots, e^{j(N-1)\phi_0}\right\} \quad (6)$$

$$H(p) = \sum_{l=0}^{L} h_l e^{-\frac{j2\pi lp}{N}} \quad (7)$$

$$\mathbf{H} = \mathrm{diag}\left\{H(0), \ldots, H(P-1)\right\}. \quad (8)$$

As long as $G \geq L$, the kth received signal after removal of CP can be formed as [5]

$$\mathbf{y}(k) = e^{j\phi_0((k-1)N+kG)}\mathbf{\Phi}_0\mathbf{W}_P\mathbf{H}\mathbf{s}(k) + \mathbf{n}(k) \quad (9)$$

where $\mathbf{n}(k)$ is an $N\times 1$ vector that represents the white Gaussian noise with zero mean and variance σ_n^2 at each sampling time. Clearly, $\mathbf{W}_P^H\mathbf{\Phi}_0\mathbf{W}_P \neq \mathbf{I}$ whenever $\phi_0 \neq 0$. Therefore, the orthogonality between subcarriers are destroyed due to the existence of nonzero CFO.

B. DIRM Algorithm

Define $\mathbf{\Phi} = \mathrm{diag}\{1, e^{j\phi}, \ldots, e^{j(N-1)\phi}\}$. In the absence of noise, it can be observed that

$$\mathbf{w}_{P+i}^H \mathbf{\Phi}^H \mathbf{y}(k) = \mathbf{w}_{P+i}^H \mathbf{\Phi}^H \mathbf{\Phi}_0 \mathbf{W}_p \mathbf{H} \mathbf{s}(k) = 0, \quad i = 1, \ldots, N-P \tag{10}$$

for $\phi = \phi_0$. Therefore

$$P(\phi) = \sum_{i=1}^{N-P} \sum_{k=1}^{K} \mathbf{w}_{P+i}^H \mathbf{\Phi}^H \mathbf{y}(k) \mathbf{y}(k)^H \mathbf{\Phi} \mathbf{w}_{P+i} \tag{11}$$

equals to zero at $\phi = \phi_0$. Based on this observation, if we define $\mathbf{Z} = \mathrm{diag}\{1, z, z^2, \ldots, z^{N-1}\}$, the cost function

$$P(z) = \sum_{i=1}^{N-P} \sum_{k=1}^{K} \mathbf{w}_{P+i}^H \mathbf{Z}^H \mathbf{y}(k) \mathbf{y}(k)^H \mathbf{Z} \mathbf{w}_{P+i} \tag{12}$$

equals to zero at $z = e^{j\phi_0}$. That is, $z = e^{j\phi_0}$ can be found from the roots of (12). However, the highly efficient polynomial rooting method cannot be directly applied on (12) because of the coexistence of both z and its complex conjugate z^*. Since the expected root $z = e^{j\phi_0}$ lies on the unit circle, it is also a root of the following polynomial:

$$\tilde{P}(z) = \sum_{i=1}^{N-P} \sum_{k=1}^{K} \mathbf{w}_{P+i}^H \mathbf{Z}^{-1} \mathbf{y}(k) \mathbf{y}(k)^H \mathbf{Z} \mathbf{w}_{P+i}. \tag{13}$$

Note that (13) is a function of z and the rooting algorithm can be applied. Since there totally exist $2N-2$ roots, denoted as $z_r, r = 1, \ldots, 2N-2$, the selection of the desired root should obey the following steps.

- Obtain $2N-2$ estimates ϕ_r as $\phi_r = \arg z_r$, $r = 1, \ldots, 2N-2$.
- The desired estimate should be chosen as the one that minimizes $P(\phi_r)$.

Remark 1: Although the roots of (12) and (13) are not totally the same, the value $z = e^{j\phi_0}$ belongs to both root categories in the absence of the noise. So we can root $\tilde{P}(z)$ instead of $P(z)$.

Remark 2: At the existence of noise, the polynomial (12) has no roots because $P(z) > 0$ for all z. Since we change z^* to z^{-1}, roots can exist for equation $\tilde{P}(z) = 0$. However, since $z = e^{j\phi_0}$ is no longer a root of $P(z)$, continuing to use the root from $\tilde{P}(z)$ will inevitably cause error in estimation. Moreover, the root obtained from of $\tilde{P}(z)$ cannot stay on the unit circle because such a root may also satisfy $P(z) = 0$, which is an impossible case. Therefore, the effect of the noise will shift the roots of $\tilde{P}(z)$ away from the unit circle. If this shift is purely on the radius direction, then the estimation results will not be changed. However, since the direction and the amount of the shift depend on the value of the noise at that specific realization, this shift is actually a random variable. Denote $d_{\min}$ as the minimum distance from all z_r to the unit circle. Table I shows the variation of the typical value of $d_{\min}$ under different SNR with parameters taken as $\Delta f = 0.4/T_s$, $P = 6$, and $N = 8$. From Table I, we can see that the roots of $\tilde{P}(z)$ are dramatically shifted away from the unit circle at low SNR. Therefore, the step from $P(z)$ to $\tilde{P}(z)$ by representing z^* as z^{-1} is actually a coarse approximation. Even for SNR as high as 20 dB, the $d_{\min}$ is still 0.03.

TABLE I
VALUE OF $d_{\min}$ FOR DIFFERENT SNR

SNR	0 dB	2 dB	4 dB	6 dB	10 dB	20 dB
d_{min}	0.2308	0.2009	0.1722	0.1454	0.1002	0.0350

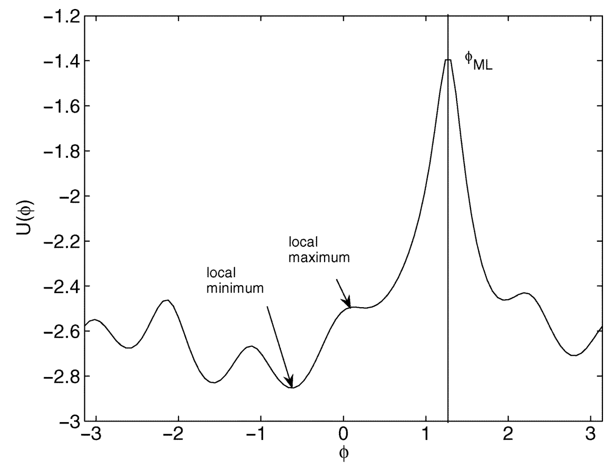

Fig. 1. Typical CFO pattern with $P = 6$, $N = 8$, $\Delta f = 0.2/Ts$, and SNR $= 10$ dB.

III. IMPROVED POLYNOMIAL ROOTING—DERM

As pointed out in [5], $e^{j\phi_0}$ is indeed a zero of (13), only in the absence of noise. Therefore, the direct rooting of (13) will inevitably cause performance degradation especially at lower SNR. In fact, the ML estimate of ϕ_0 is proved to be [5], [6]

$$\phi_{ML} = \arg\min_{\phi} P(\phi). \tag{14}$$

Based on this observation, an adaptive approach that can achieve the local minimum is proposed in [5]. However, there are several shortcomings of using the adaptive approach. One is well known to be the slow convergence rate and difficulty in choosing an appropriate step size. The other is the requirement of a "good enough" initial point to ensure the convergence at the global minimum point. The latter is even critical since the region for a good initial point is usually small. A typical CFO pattern, defined as the plot for $U(\phi) = \log(1/P(\phi))$, is shown in Fig. 1 with parameters $P = 6$, $N = 8$, $\Delta f = 0.2/Ts$, and SNR $=$ 10 dB. For this example, if the initial estimate of Δf stays outside the region $[0.05/T_s, 0.3/T_s]$, the adaptive algorithm will give a false estimate.

Another way to achieve an ML detection of CFO is suggested in [3] and [4], where the value of ϕ_0 can be found by evaluating $P(z)$ on the unit circle $z = e^{j\phi}$ at all possible values of $\phi \in [-\pi, \pi]$. Although this searching method yields the ML estimate, it is computationally quite expensive. Moreover, the complexity and the estimation accuracy strictly depend on the grid that is used during the search.

Inspired by all these facts, we propose an improved polynomial rooting approach that yields the ML estimate and is, meanwhile, quite computationally efficient. This approach, as can be seen from literature, has never been proposed in [3] and [4] and other blind CFO estimation works.

First, we list the mathematical rule to find the global minimum point for a cost function.

- Obtain the solutions for all local minimum/maximum as well as the global minimum/maximum by letting the derivative of the cost function be zero.
- Put these solutions back to the original cost function and select the minimum after comparisons.

Based on this rule, we define

$$\begin{aligned} Q(\phi) &= \frac{\partial P(\phi)}{\partial \phi} \\ &= \sum_{i=1}^{N-P} \sum_{k=1}^{K} \mathbf{w}_{P+i}^{H} \mathbf{\Phi}^{H} \\ &\quad \times \left(\mathbf{D}\mathbf{y}(k)\mathbf{y}(k)^{H} - \mathbf{y}(k)\mathbf{y}(k)^{H}\mathbf{D} \right) \mathbf{\Phi}\mathbf{w}_{P+i} \end{aligned} \quad (15)$$

where $\mathbf{D} = \text{diag}\{0, 1, 2, \ldots, N-1\}$. Obviously, one of the roots to (15) must be ϕ_{ML} and others are local minimum/maximum of $P(\phi)$. This holds whether or not there is noise.

Replacing $\mathbf{\Phi}$ by $\mathbf{Z}$, we know $z = e^{j\phi_{ML}}$ is one of the roots for

$$\begin{aligned} Q(z) = \sum_{i=1}^{N-P} \sum_{k=1}^{K} & \mathbf{w}_{P+i}^{H} \mathbf{Z}^{H} \\ & \times \left(\mathbf{D}\mathbf{y}(k)\mathbf{y}(k)^{H} - \mathbf{y}(k)\mathbf{y}(k)^{H}\mathbf{D} \right) \mathbf{Z}\mathbf{w}_{P+i}. \end{aligned} \quad (16)$$

Since $z = e^{j\phi_{ML}}$ has a unit norm, we can replace $\mathbf{Z}^{H}$ with $\mathbf{Z}^{-1}$, and $z = e^{j\phi_{ML}}$ can also be found from one of the roots of

$$\tilde{Q}(z) = \sum_{m=-N+1}^{N-1} a_m z^m \quad (17)$$

where

$$a_m = \sum_{q-p=m} [\mathbf{T}]_{pq} \quad (18)$$

$$\mathbf{T} = \Im \left\{ \sum_{i=1}^{N-P} \mathbf{W}_{P+i}^{H} \mathbf{D} \mathbf{R}_{\mathbf{y}} \mathbf{W}_{P+i} \right\} \quad (19)$$

$$\mathbf{R}_{\mathbf{y}} = \frac{1}{K} \sum_{k=1}^{K} \mathbf{y}(k)\mathbf{y}(k)^{H} \quad (20)$$

$$\mathbf{W}_{P+i} = \text{diag}\{\mathbf{w}_{P+i}\} \quad (21)$$

with $\Im\{\cdot\}$ representing the imaginary part of the inside function.

Since totally $2N-2$ roots exist for $\tilde{Q}(z)$ (again, call them as $z_r, r = 1, \ldots, 2N-2$), the desired root $z = e^{j\phi_{ML}}$ should be selected according to two criteria.

- Choose roots z_{r_t}, $t = 1, \ldots, T$, $r_t \in \{1, \ldots, 2N-2\}$ that stay on the unit circle, where T is an integer smaller than or equal to $2N-2$.
- Obtain ϕ_{r_t} as $\phi_{r_t} = \arg z_{r_t}$. The desired estimate ϕ_{ML} is the one that minimizes $P(\phi_{r_t})$ among all ϕ_{r_t}.

Remark 3: The root $z = e^{j\phi_{ML}}$ always stays on the unit circle, and other ϕ_{rt} are the local minimum/maximum, global maximum for $P(\phi)$. This is quite different from DIRM where the roots of $\tilde{P}(z)$ are shifted away from the unit circle by the noise.

Remark 4: Since the item $z = e^{j\phi}$ is involved in $\tilde{Q}(z)$, the DERM algorithm can achieve a CFO estimation region as $[-\pi, \pi)$. Note that this is the same region for CFO estimation in DIRM [3], [4] and is also the maximum possible region for all CFO estimation methods.

Remark 5: Since the highest order of both polynomials $\tilde{P}(z)$ and $\tilde{Q}(z)$ is $2N-2$, the complexities of both rooting approaches are considered to be the same and can be approximated as $O(N^3)$. Since $T \leq 2N-2$, the number of comparison involved in DERM may be slightly smaller than that in DIRM. However, since the complexity of the comparison is much smaller compared to the rooting approach, the complexities of both rooting algorithms can be considered as the same.

Remark 6: The identifiability problem is not considered here since it is an independent issue. Actually, to guarantee the estimate of ϕ_0, one could use the distinct spaced virtual carrier or the carrier hopping proposed in [5]. Then a different cost function of ϕ, or equivalently $e^{j\phi}$, could be obtained where the polynomial rooting can be applied, again, by replacing $e^{j\phi}$ with z. One should only make sure that the first-order derivative of the cost function is rooted instead of a direct rooting of the cost function itself.

IV. Numerical Results

We provide several simulations in this section to validate the proposed theoretical analysis. For all numerical examples, a three-ray channel model is used with exponential power delay profile given by

$$E\left\{|h(l)|^2\right\} = \exp\left(-\frac{l}{10}\right), \qquad l = 0, 1, 2. \quad (22)$$

The phase of each channel ray is uniformly distributed over $[0, 2\pi)$. The data transmitted are modulated by quadrature phase-shift keying (QPSK), and the normalized CFO is taken as large as $\phi_0 = 0.8\pi$. All results are averaged over $M = 500$ Monte Carlo runs. The normalized estimation mean-square error (NMSE) is defined as

$$\text{NMSE} = \frac{1}{M} \sum_{i=1}^{M} \frac{(\hat{\phi}_i - \phi_0)^2}{\phi_0^2} \quad (23)$$

where the subscript i refers to the ith simulation run.

In the first example, we compare the DERM with the searching-based ML estimation. The parameters are taken as $P = 6$, $N = 8$, and $K = 10$. The grid size for ML searching is taken as 0.04 and 0.001, respectively, over the entire region $[-\pi, \pi]$. Therefore, the resolution of these two searching approaches may be expressed as $|\hat{\phi} - \phi_{ML}| < 0.02$ and $|\hat{\phi} - \phi_{ML}| < 0.0005$, respectively. The NMSEs of CFO estimation versus SNR results are shown in Fig. 2. It can be seen that the DERM gives exactly the same performance as that of ML searching with the grid size 0.001. This is quite reasonable since the proposed DERM is also an ML estimator for CFO. However for the grid size of 0.04, the searching-based ML algorithm performs worse than the proposed DERM and reaches a lower bound after SNR $=$ 12 dB. This is a direct result of its lower resolution. Therefore, the performance of the

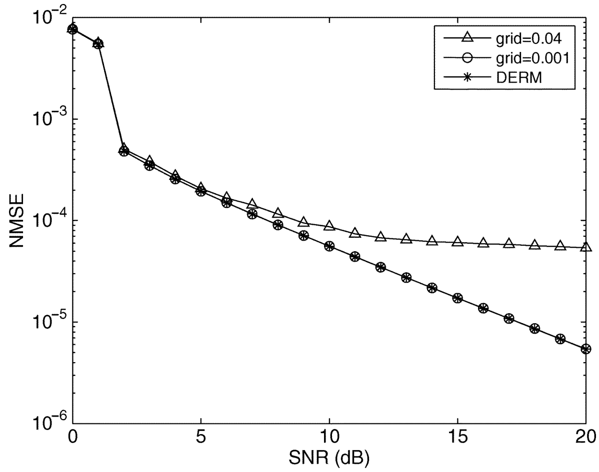

Fig. 2. DERM versus searching-based ML estimators.

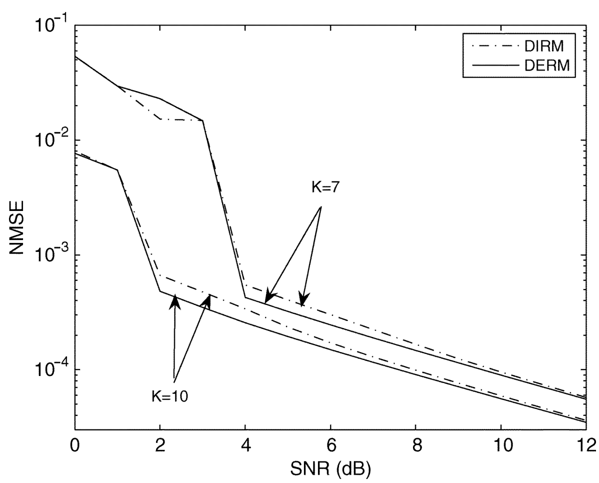

Fig. 3. DERM versus DIRM for different values of K.

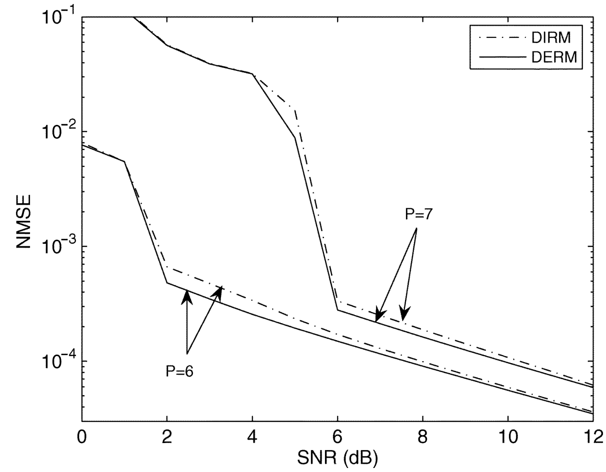

Fig. 4. DERM versus DIRM for different values of P.

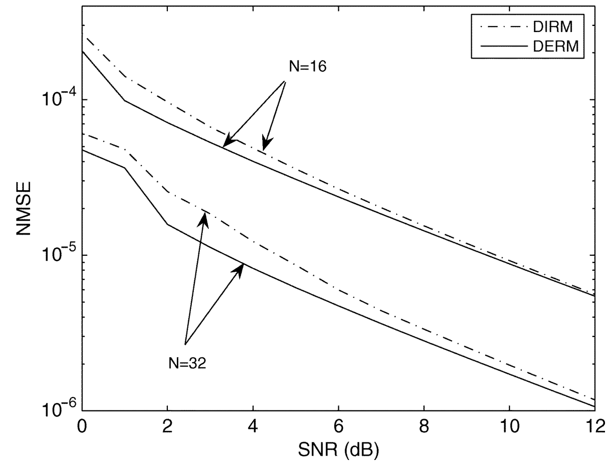

Fig. 5. DERM versus DIRM with different N but the same P/N.

ML searching method is crucially related to the grid size. However, reducing the grid size will greatly increase the complexity of the searching-based method.

The performance of the DIRM and the DERM is compared under various scenarios. First, we take $P = 6$, $N = 8$ and examine the performance of both rooting methods by changing the value of K. The NMSEs of CFO estimation versus SNR results are shown in Fig. 3. Next, we fix $K = 10$, $N = 8$ and compare the performance of both rooting methods by changing the value of P. The total signal power in each block is the same for different P. The NMSEs versus SNR results are displayed in Fig. 4. Finally, we fix $K = 10$, $P/N = 3/4$ and compare the two algorithms by varying the value of N. The NMSEs versus SNR results are displayed in Fig. 5.

Clearly, from all numerical examples, we see that the DERM outperforms the DIRM at all SNR regions. The performance of the DIRM can only approaches that of DERM asymptotically at high SNR. This is because the proposed DERM is exactly the ML estimator, but the DIRM is only suboptimal.

V. Conclusion

In this letter, we proposed a blind search-free CFO estimator for OFDM systems by exploiting the polynomial rooting method. It is pointed out that the proposed search-free method is also an ML estimator for CFO estimation, and it outperforms the existing search-free technique that ignores the effect of the noise. Simulation results clearly show the performance improvement of the proposed method.

References

[1] W. Y. Zou and Y. Wu, "COFDM: An overview," *IEEE Trans. Broadcast.*, vol. 41, no. 1, pp. 1–8, Mar. 1995.

[2] T. Pollet, M. van Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Trans. Commun.*, vol. 43, no. 2, pp. 191–193, Feb. 1995.

[3] H. Liu and U. Tureli, "A high-efficiency carrier estimator for OFDM communications," *IEEE Commun. Lett.*, vol. 2, no. 4, pp. 104–106, Apr. 1998.

[4] U. Tureli, D. Kivanc, and H. Liu, "Experimental and analytical studies on a high-resolution OFDM carrier frequency offset estimator," *IEEE Trans. Veh. Technol.*, vol. 50, no. 2, pp. 629–643, Mar. 2001.

[5] X. Ma, C. Tepedelenlioglu, G. B. Giannakis, and S. Barbarossa, "Non-data-aided carrier offset estimators for OFDM with null subcarriers: Identifiability, algorithms, and performance," *IEEE. Trans. Commun.*, vol. 19, no. 12, pp. 2504–2515, Dec. 2001.

[6] B. Chen, "Maximum likelihood estimation of OFDM carrier frequency offset," *IEEE Signal Process. Lett.*, vol. 9, no. 4, pp. 123–126, Apr. 2002.