Hybrid Cramer-Rao Bound on Carrier and Sampling Frequency Offset Estimation for OFDM Systems in Rayleigh Fading Channels

Qi Cheng1 · Mosin (Andre) Russell1

Accepted: 28 August 2022 / Published online: 29 September 2022 © The Author(s) 2022

Abstract
For the carrier frequency offset (CFO) and sampling (clock) frequency offset (SFO) estimation, the hybrid Cramer-Rao bound (HCRB) is developed when the CFO, SFO, information-bearing symbols are deterministic and channel coefficients are random. Both noise and channel coefficients are complex Gaussian. The HCRB is a lower bound on the mean squared estimation error for any unbiased estimator of a parameter. For the HCRB to be applicable, it is necessary for deterministic parameters to be identifiable (uniquely determined). Some necessary identifiability conditions of some deterministic parameters are found and presented. The HCRB is dependent on the initial time instant. The HCRB is used to assess the performances of some existing methods via simulation. Our results demonstrate that even the best performance is still around 10 dB higher than the HCRB. Further effort is needed to develop more accurate methods.

Keywords OFDM · CFO · SFO · Hybrid Cramer-Rao bound

1 Introduction
Orthogonal frequency division multiplexing (OFDM) is a multicarrier modulation/demodulation scheme which has been adopted by the standards for digital audio/video broadcasting and WLAN. OFDM systems improve spectral efficiency but are more sensitive to carrier frequency offset (CFO) and sampling clock frequency offset (SFO). The CFO and SFO are mainly caused by the mismatch between the oscillators of the transmitter and the receiver in OFDM systems. They destroy the orthogonality between subcarriers and create inter-carrier interference after demodulation. They have to be estimated from measurement and compensated.

Qi Cheng
q.cheng@westernsydney.edu.au

Mosin (Andre) Russell
a.russell@westernsydney.edu.au

1 School of Engineering, Design and Built Environment, Western Sydney University, Locked Bag 1797, Perth MC SW 1797, Penrith, NSW, Australia
For the joint estimation of the CFO and SFO, most methods are built on the correlation of the demodulated measurement at the same pilot carrier in two data symbols. The demodulated measurement is obtained by passing a data symbol through the DFT. A correlation can be calculated from two consecutive data symbols or from two nonconsecutive data symbols. In this paper, the former will be called single-lag correlation and the latter multiple-lag correlation. When the CFO and SFO are both small enough, the phase of the single-lag correlation for a pilot subcarrier can be approximated by a linear function of the CFO and SFO.

The method in [1] determines those phases for all pilot subcarriers using single-lag correlations only. The method in [2] obtains those phases using products of single-lag and multi-lag correlations. The method in [3] finds those phases via the rooting of a polynomial. That polynomial is constructed from the Taylor series expansion of an exponential function based on single-lag and multi-lag correlations. Next, in [1–3], optimally weighted least squares solutions of the CFO and SFO, are obtained from the same system of linear equations. Their optimum weighting matrices turn out to be equal to the same diagonal matrix, where each diagonal entry is proportional to the squared channel frequency response at a pilot subcarrier.

The estimates of the CFO and SFO given in [4], are determined from a least squares formulation of the demodulated measurements of the two long symbols in the WALN preamble. This method requires a two-dimensional exhaustive search. The authors of [5] use a second-order Taylor series approximation to obtain closed-form expressions of the CFO and SFO. In particular, the SFO is determined from equation (17) of that paper, in which products of single-lag correlations at different pilot subcarriers, are used. From an SFO estimate, a CFO estimate can be directly calculated from (12) of the same paper. The method in [4] is a maximum likelihood algorithm when channel coefficients are deterministic. However, for random channel coefficients in this paper, this is no longer the case, and hence the method in [4] is not guaranteed to provide the optimum performance.

In [6], a different approach is proposed to obtain the CFO estimate. This approach first adds the single-lag correlations at all pilot subcarriers and then calculates the single phase of the sum. The CFO estimate is next obtained from this phase. Afterwards, this CFO estimate is subtracted from the estimated phases of single-lag correlations, and the SFO estimate is successively obtained from the obtained phase differences via a weighted procedure.

The method in [7] exploits the structure of the long symbol cyclic prefix and the two long symbols in the WLAN preamble. Thus it cannot be directly applied to demodulated measurement at pilot subcarriers of data symbols.

Cyclostationarity is used to develop an SFO estimation method in [8]. It does not require pilot subcarriers, and hence improves spectral efficiency. But to obtain reliable cyclostationary statistics, a much larger number of data symbols are required. The method in [8] does not consider the estimation of the CFO. Special structures of signals in digital radio mondiale broadcasting systems are exploited to develop the methods in [9, 10]. But they are not applicable to general OFDM systems. Hence, the methods in [8–10] will not be considered in simulation for comparison.

For later referencing convenience, the existing methods [1–3, 5, 6], will be abbreviated as: the single-lag correlation (SLC) method [1], the multi-lag correlations (MLC) method [3], the multi-lag correlation product (MLCP) method [2], single-lag correlation product (SLCP) method [5]. The method in [6] will be called successive interference cancellation.

The mean-squared estimation errors of the CFO and SFO are an important factor to compare the performances of various methods. These two errors of any unbiased
method are lower bounded by a limit, called the Cramer-Rao bound (CRB). The CRB indicates the room for performance improvement. The conventional CRB treats deterministic variables as unknown parameters and random variables (such as the amplitude of an incoming wave and noise quantities) as nuisance parameters. The HCRB is the extended Cramer-Rao bound when some unknown parameters (such as channel coefficients) are random. The HCRB has not been used in performance assessment for the joint estimation of the CFO and SFO in the literature. In [2–5, 11], the CRB was derived where the channel coefficients were chosen to be deterministic (fixed). That CRB should be used for performance assessment, only in the deterministic channel case, as in Figures 1–2 of [3]. However, that CRB was also used in simulation for random (different) channels, in [2, 4, 5, 11]. Due to a lack of details in those papers, it is not clear how that CRB was calculated for random channels. The deterministic channel case is not realistic because in practice, channel varies from time to time. Thus it is necessary to consider the HCRB.

In addition to the HCRB, two other bounds were also proposed: modified CRB (MCRB) in [12] and the Miller-Chang bound in [13]. Given various bounds, a crucial issue is which one is achievable under the same set of conditions by some methods. As explained in the summary and discussion section of [14], the MCRB is always no larger than the HCRB. This means that under the regularity conditions mentioned in [14], the MCRB is not achievable and should not be used. To derive the Miller-Chang bound, the covariance matrix of the first-order partial derivatives is first derived based on the probability density function of the measurement conditioned on random parameters and then inverted to yield the conditional CRB matrix; the expectation is next applied to the conditional CRB matrix using the prior probability density function of random parameters. For the problem in this paper, the conditional CRB matrix contains channel coefficients in denominators, and thus taking expectation does not lead to a closed form expression. Therefore, it is impossible to identify achievability conditions for the Miller-Chang bound. In [14], it is proven that the HCRB can be achieved by the maximum likelihood/maximum a posteriori (ML/MAP) estimator (p. 12 of [15]) for unbiased estimates, under certain regularity conditions, i.e., the HCRB is the currently known tightest bound. Hence, in this paper, only the HCRB is used to assess the estimation performance.

The implementation of the ML/MAP estimator requires a highly multi-dimensional search. Its execution is very time-consuming and its performance also depends how close the initial estimates are located to the true values. All these issues are worth further investigation. Thus in this paper, this estimator is not included in comparison.

In this paper, the HCRBs for the CFO and SFO are derived for random channel coefficients. Some properties of the HCRBs are discovered. Deterministic parameters, such as the CFO, SFO, and information-bearing symbols, should be uniquely determined. Otherwise, the estimates of those parameters may be biased and then the HCRB is not guaranteed to be the lower-bound of their mean-squared errors. Some necessary identifiability conditions for them are found and presented in this paper. These conditions can also be used as a guide in practice to ensure the identifiability of those parameters.

As described before, in most methods, the phases of demodulated measurement is used to determine estimates of the CFO and SFO. The estimates can be accurate only when both the CFO and SFO are sufficiently small. The SFO is already very small, but the CFO is not. In this paper, it is within one subcarrier spacing, represented by a number in the range (−0.5, 0.5]. The CFO value for those method should be tenths or hundredths of one subcarrier spacing, such as 0.01 in [1], 0.02 used in [2, 3, 6]. Due to this reason, one has to determine a coarse estimate of the CFO first, and then perform a CFO compensation to reduce its value.
remaining value of the CFO is called the residual CFO (RCFO). One can refer to [7] on the coarse CFO estimation.

The following notations will be used throughout the paper—: conjugation, $^T$: transpose, $^H$: conjugate transpose; $I_n$: an $n \times n$ identity matrix; $0_{n_1 \times n_2}$: an $n_1 \times n_2$ matrix (including vector as a special case) with all elements equal to 0; $| \cdot |$: absolute value; $\Re \{ \cdot \}$: real part of a complex number; $\Im \{ \cdot \}$: imaginary part, of a complex number; $\| \cdot \|$: 2-norm of a vector; $E \{ \cdot \}$: the statistical expectation taken with respect to random variables involved; diag: a diagonal or block diagonal matrix.

The organization of this paper is as follows. In Sect. 2, the system model is described and assumptions on noise and channel coefficients are given. Identifiability conditions are presented in Sect. 3. The HCRB is presented and its properties are discussed in Sect. 4. Section 5 describes simulation setup, presents simulation investigation on the impact of the initial time instant and comparison of three existing methods against the HCRB. Section 6 concludes the paper.

2 System Model

An OFDM system, with $N$ subcarriers, a transmitter and a receiver, is considered. The $N$ subcarriers have the index set $[-(N/2 - 1), \ldots, -1, 0, 1, \ldots, N/2]$. Among them, $P$ subcarriers are used to transmit information-bearing symbols. Without loss of generality, $P$ is assumed to be an even number, as defined by standards, such as that on p. 247 of [16]. Those subcarriers are called active subcarriers, with the index set

$$\mathcal{P} = [-P/2, \ldots, -1, 1, \ldots, P/2].$$

(1)

In this system, $I$ data symbols are received. For the $i$-th data symbol, $P$ information-bearing symbols and pilot symbols

$$s_{i,-P/2}, \ldots, s_{i,-1}, s_{i,1}, \ldots, s_{i,P/2}$$

(2)

are first modulated onto active subcarriers at the transmitter, and next transmitted through a frequency selective channel. Pilot symbols are used for the purpose of channel estimation and frequency synchronization. In current standards, pilot symbols are chosen to be equal to certain known values (such as p. 290 of [16]). In this paper, we consider $K$ pilot symbols in each data symbol, with their subcarrier indices denoted by

$$p_1, p_2, \ldots, p_K.$$  

(3)

The channel is modeled as an $L$-tap finite impulse response system with complex coefficients $h_l$ for $l = 0, 1, \ldots, L - 1$.

The CFO is denoted by $\psi$ and the SFO by $\zeta$. They are the normalized values, i.e., actual offsets divided by the subcarrier spacing and sampling frequency respectively. $\psi$ is assumed to be strictly within one subcarrier spacing, i.e.,

$$-1/2 < \psi \leq 1/2.$$  

(4)

The SFO is a very small number, i.e.,

$$|\zeta| \leq \Delta_\zeta \ll 1$$

(5)

where $\Delta_\zeta$ is a positive number satisfying
Equations (4) and (5) can be met by some standard-compliant systems, such as the 802.11ac-based ones. In the 802.11ac standard [16], the carrier frequency is 2.4 GHz, the sampling frequency is 20 MHz, and transmit/receive oscillators’ precision tolerance is specified to be less than ±25 ppm, that results in the CFO and SFO (between the transmitter and receiver oscillators) in the range from −50 to 50 ppm. Hence, the CFO lies in the range $[-50, 50]$ ppm (50-ppm×2.4 GHz) and the SFO in the range $[-1, 1]$ kHz (50-ppm×20 MHz). In this case, the subcarrier spacing is 312.5 kHz. Thus

$$|\psi| \leq 120/312.5 = 0.384 \leq 1/2, \ |\xi| \leq 1 \text{kHz}/20 \text{MHz} = 0.00005.$$  

(7)

In the same standard, $N = 64$, $P = 56$, then for the value of $\Delta_\zeta$ in (7), the assumption (6) is also met.

Equations (4), (5) and (6) are used to prove (36), which is crucial to the unique determination of the CFO and SFO from a data covariance matrix in Sect. 3.2.

Let

$$\mathbf{F} = [f_0, f_1, \ldots, f_{L-1}]$$

(8)

$$f_i = [z_0^{-(iP/2)}, \ldots, z_{0}^{j}, z_0, \ldots, z_0^{(iP/2)}]^T, \ z_0 = e^{j2\pi/N}$$

(9)

$$\mathbf{h} = [h_0, h_1, \ldots, h_{L-1}]^T$$

(10)

$$\mathbf{Fh} = [H_{-P/2}, \ldots, H_{-1}, H_1, \ldots, H_{P/2}]^T.$$  

(11)

Denote by $n_{0,i}$ the time instant of the first sample of the $i$-th data symbol (after the cyclic prefix), and by $N_{cp}$ the length of the cyclic prefix. Then, according to [17], after dropping the cyclic prefix, the measurement of $n$-th sample of the $i$-th data symbol at the receiver is given by

$$x_i(n) = z_0^{n_0,i+N_{cp}+n(1+\zeta)}, \sum_{k \in P} z_0^{k(n+1+\zeta)} H_k s_{i,k} / \sqrt{N}$$

(12)

In (12), $n_{0,i}$ has an impact on the HCRB values. The performance of a least-squares method based on (12) may be affected by the value of $n_{0,i}$; $n_{0,i}$ accounts for the combined impact of the initial time of transmission and the relative propagation delay of the transmitted signal through a multipath channel. Its impact will be investigated in the numerical study section. In this paper, the initial time instant is chosen as the first sample of the cyclic prefix of the $I_{ini}$-th data symbol. Then

$$n_{0,i} = (i - I_{ini})(N + N_{cp}).$$  

(13)

Define

$$\mathbf{x}_i = [x_i(0), x_i(1), \ldots, x_i(N-1)]^T$$

(14)
\[
E = \text{diag}[1, z_0^{(1+\zeta)\psi}, \ldots, z_0^{(N-1)(1+\zeta)\psi}]
\]

\[
v_p = [1, z_0^{p(1+\zeta)}, \ldots, z_0^{(N-1)p(1+\zeta)}]^T / \sqrt{N}, \ p \in \mathcal{P}
\]

\[
V = [v_{-p/2}, \ldots, v_{-1}, v_1, \ldots, v_{p/2}]
\]

\[
\Omega_i = \text{diag}[z_0^{-n_{0i}(P/2)\zeta}, \ldots, z_0^{-n_{0i}P\zeta}, z_0^{n_{0i}P}, \ldots, z_0^{n_{0i}(P/2)\zeta}]
\]

\[
S_l = \text{diag}[s_{i,-p/2}, \ldots, s_{i,-1}, s_{i,1}, \ldots, s_{i,p/2}].
\]

Then the vector representation of the \(i\)-th received data symbol can be written as

\[
x_i = z_0^{(n_{0i}+N_p)(1+\zeta)\psi} EV\Omega_i S_l F h_i.
\]

Define \(y_i\) to be the noise corrupted \(\chi_i\):

\[
y_i = x_i + n_i, \quad i = 1, \ldots, I
\]

where \(n_i\) is the complex measurement noise vector of the \(i\)-th data symbol.

In most practical systems, information-bearing symbols \([s_i(p), \ \forall p \in \mathcal{P} - p_j]_{j=1}^K\) are taken from one of the BPSK, QPSK and QAM constellations, etc. Thus they only have one or a few discrete values. But in most existing methods, such as those in [1–3, 5, 6, 17], the discrete-value feature of information-bearing symbols is not exploited. The methods in [1–3, 5, 6] will be compared against the HCRB in the simulation section. Hence, in the derivation of the HCRB, they are assumed to be complex analogue quantities.

Exactly speaking, each \(n_i\) in (21) is a realization of the random vector \(N_i\). \(N_i\) and \(N_j\) are independent for \(i \neq j\). Each \(N_i\) contains independent and identically distributed (i.i.d) complex Gaussian variables with mean zero and variance \(\sigma^2\). The real part and imaginary part of each complex noise variable are also i.i.d.. Complex channel coefficients \(h_i\), \(l = 0, 1, \ldots, L - 1\) in (20), are also realizations of random variables \(H_i\), which are i.i.d. complex Gaussian variables with mean zero and variance \(\sigma_p\). The real part and imaginary part of each channel coefficient are i.i.d. as well. Noise variables are independent of channel coefficients. The above assumptions can be represented in the formulas below:

\[
E\{n_i n_j^H\} = \sigma_i n_{ij} I_{N}, \quad E\{n_i n_j^T\} = 0_{N \times N}
\]

\[
E\{hh^H\} = \text{diag}[\sigma_0, \sigma_1, \ldots, \sigma_{L-1}] \overset{\text{def}}{=} D_{\sigma_k}, \quad E\{hh^T\} = 0_{L \times L}
\]

\[
E\{hn_i^H\} = 0_{L \times N}, \quad E\{hn_i^T\} = 0_{L \times N}
\]

where \(\delta_{ij}\) is the Dirac function which is equal to one only when \(i = j\) and zero when \(i \neq j\). The third-order covariances involving channel coefficients and noise quantities are zero due to zero-mean Gaussianity assumption:

\[
E\{n_{i_1}(k_1)n_{i_2}(k_2)(n_{i_3}(k_3))^*\} = 0, \quad E\{h_{i_1}h_{i_2}(h_{i_3})^*\} = 0
\]
∀i₁, i₂, i₃ ∈ [1, I], ∀k₁, k₂, k₃ ∈ ℙ, ∀l₁, l₂, l₃ ∈ [0, L − 1], where nᵢ(k) is the k-th element of nᵢ. The fourth-order variances involving channel coefficients and noise quantities are given below, also due to Gaussianity assumption:

\[ E[|nᵢ(k)|^4] = 2\sigma^2, \quad E[|hᵢ|^4] = 2(\sigma_i)^2. \]  \hspace{1cm} (27)

The most common problem in the literature is to estimate the CFO, SFO, unknown information-bearing symbols, \( \sigma, \sigma_i, l = 0, \ldots, L - 1 \) and unknown one realization of channel coefficients, from the noisy received data \( y^T_i \) in (21). In this paper, identifiability conditions of the CFO, SFO, information-bearing symbols, and one realization of the random channel will be investigated, and an explicit expression of the HCRB will be derived to assess the performance of any CFO and SFO estimator. The identifiability is defined as the unique determination of the those parameters, from the noise-free version of the data symbols in (12).

### 3 Identifiability Conditions

The data symbol in (12) can be regarded as the measurement for one realization of the random channel. If some of those parameters cannot be uniquely determined, their estimates may not be unbiased and the HCRB may not be applicable.

To the authors’ knowledge, the identifiability has not been discussed in the literature, including [1–3, 5, 6]. Hence, in this section, identifiability conditions will be studied and presented.

In Sect 3.1, the full column rank conditions of a measurement covariance matrix are derived. Those conditions are then used in Sect. 3.2 to develop a method, which can provide unique solutions of deterministic parameters and one realization of channel coefficients.

### 3.1 Multiple-Symbol Based Covariance Matrix and its Rank Conditions

Let \( b = [b_0, b_1, \ldots, b_P]^T \) be a complex vector such that the polynomial

\[ B(z) = b_0^*z^P + b_1^*z^{P-1} + \cdots + b_{P-1}^*z + b_P^* \]  \hspace{1cm} (28)

have roots \( z_0^{(p+\psi)(1+\zeta)} \), \( \forall p \in ℙ \). With no loss of generality, it is assumed that \( b_0 \neq 0 \). Construct the following \( N \times (N - P) \) matrix

\[ \mathbf{B} = \begin{bmatrix}
    b_0 & b_1 & \cdots & b_P & 0 & \cdots & 0 \\
    0 & b_0 & b_1 & \cdots & b_P & \ddots & \vdots \\
    \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
    \vdots & \cdots & \ddots & \ddots & \ddots & \ddots & \ddots \\
    0 & \cdots & \cdots & 0 & b_0 & b_1 & \cdots & b_P
\end{bmatrix}^T. \]  \hspace{1cm} (29)

Then

\[ x_i^H\mathbf{B} = 0_{1 \times (N - P)}, \quad i = 1, \ldots, I. \]  \hspace{1cm} (30)

Let
be a \((P + 1)\)-dimensional subvector of \(x_i\) for \(k = 1, 2, \ldots, N - P\), and define
\[
X_i = [x_i^{(1)}, x_i^{(2)}, \ldots, x_i^{(N-P)}].
\]
Then (30) can be rewritten as
\[
X_i^H b = 0_{(N-P) \times 1}, \quad i = 1, \ldots, I \rightarrow \left( \sum_{i=1}^{I} X_i X_i^H \right) b = 0. \tag{32}
\]

Next, the minimum value \(I_0\) on \(I\) and other conditions for the unique solution of \(\psi(1 + \zeta)\) from \(b\) will be investigated.

To this purpose, the following \((P + 1) \times (P + 1)\) covariance matrix
\[
\sum_{i=1}^{I_0} X_i X_i^H \tag{33}
\]
should have rank \(P\). According to (74), this rank condition in turn requires
\[
\sum_{i=1}^{I_0} (\Omega_i s_i) \Phi \Phi^H (\Omega_i s_i)^* \tag{34}
\]
to possess full rank \(P\).

**Theorem 1** For the matrix in (34) to be of full rank \(P\), it is necessary that (i)
\[
I_0 = \left\lfloor \frac{P}{N - P} \right\rfloor \tag{35}
\]
and (ii) the matrix \([\Omega_1 s_1, \Omega_2 s_2, \ldots, \Omega_{I_0} s_{I_0}]\) has full column rank, where \(\lfloor \cdot \rfloor\) denotes the ceiling of a number.

**Proof** See Appendix A. \(\square\)

The first condition (35) in Theorem 1 shows a minimum value on the number of symbols required for the full rank condition of the matrix in (34). The second condition in Theorem 1 depends on information-bearing symbols, and the SFO. For the parameters used in the simulation section, via numerical calculation, the condition is satisfied when \(s_i, i = 1, \ldots, I_0\) are independent.

**Theorem 2** When (i) the two conditions in Theorem 1 are satisfied and (ii) \(H_p \neq 0 \forall p \in \mathcal{P}\), the matrix in (33) is of rank \(P\).

**Proof** See Appendix B. \(\square\)

The condition of \(H_p \neq 0 \forall p \in \mathcal{P}\) is not restrictive, because in practice, all active subcarriers are used to transmit information-bearing symbols and pilot symbols, hence from (12), their frequency responses can not be zero. Otherwise, the information-bearing symbols transmitted over those active subcarriers are lost and frequency responses for pilot symbols can not be obtained, at the receiver end.
3.2 Unique Identification

Since the matrix in (33) is of rank $P$, thus there exists only one eigenvector corresponding to the eigenvalue 0. This eigenvector is unique up to a complex scalar, and thus the $P$ roots of $B(z)$ (defined in (28)): $z_0 = e^{(p+\psi)(1+\zeta)}$, $\forall p \in P$ can be uniquely determined. Using (4), (5) and (6), one knows that

$$|\psi p (1 + \zeta)| \leq |\psi p | |1 + \zeta| < (P/2 + 1/2)(1 + \Delta \zeta) < N/2,$$

(36)

and hence the phases of all those roots fall within the range ($-\pi, \pi$) and can be found without ambiguity. Denote those $P$ phases by $\alpha_p, p \in P$. By adding them, one can find the value of $\psi p (1 + \zeta)$ as

$$\psi p (1 + \zeta) = \sum_{\forall p \in P} \alpha_p / P.$$  

(37)

Next, from (74), one can construct the following compensated covariance matrix

$$E_0^* \sum_{i=1}^{I_0} X_iX_i^H \cdot E_0 = V_0 \cdot \mathcal{H} L_0^0 \bigg( L_0^{(1+\zeta)} \bigg) \cdot \Phi \Phi^H (\Omega_lS_l)^* \mathcal{H}^* \cdot V_0^H.$$  

(38)

The matrix in (38) also has rank $P$. Denote by $U$ the set of unitary eigenvectors corresponding to $P$ positive eigenvalues, and let $U_L$ and $U_H$ be the submatrices of $U$, comprising of the top $P$ rows and the bottom $P$ rows respectively. Then the matrix $(U_L^H U_L)^{-1} U_L^H U_H$ contains the following eigenvalues $e^{J^p (1+\zeta)}$, $\forall p \in P$. Since $P < N$, thus those phases also fall within the range ($-\pi, \pi$). Arrange those $P$ phases in increasing order and denote those ordered-phases by $\beta_p, p \in P$. Then, one can find the unique value of $\zeta$ as

$$\zeta = \left( \sum_{\forall p \in P} \beta_p / p \right) / P - 1.$$  

(39)

Plugging (39) into (37) gives a unique solution of $\psi$

$$\psi = \left( \sum_{\forall p \in P} \alpha_p / \left( \sum_{\forall p \in P} \beta_p / p \right) \right).$$  

(40)

When $S_1$ is known, then from (42) and the value of the SFO in (39), one can obtain unique channel coefficients as

$$h = z_0^{-n_0 + N_p (1+\zeta)\psi} (F^H F)^{-1} F^H \Omega_1^* S_1^* (V^H V)^{-1} V^H E^* x_1.$$  

(41)

If one forms the following demodulated vector

$$x_i' = z_0^{-n_0 + N_p (1+\zeta)\psi} (\Omega_i)^* (V^H V)^{-1} V^H E^* x_i = S_i F h,$$  

(42)

then $s_i', i = 2, \ldots, I$ can be uniquely found.

As mentioned in [7], two identical data symbols are required to determine a coarse CFO. Hence, in the received data (20), the first two data symbols are assumed to be identical and their information-bearing symbols are assumed to be known, and the total number of data symbols is no less than $I_0 + 1$, where at least the 2nd, 3rd, ..., $(I_0 + 1)$-st
data symbols contain independent information-bearing symbol vectors, to meet the conditions required by Theorem 1.

4 Hybrid Cramer-Rao Bound

In this section, the HCRB will be derived and its properties will be discussed.

Define the selection matrix $\mathbf{J}$ to be a $(P - K) \times P$ submatrix of $\mathbf{I}_P$ with the pilot-symbol rows deleted. Then the unknown information bearing symbols are only contained in the vector $\mathbf{J}_s$, where $s_i = [s_{i-P/2}, \ldots, s_{i-1}, s_i, s_{i+1}, \ldots, s_{i+P/2}]^T$. Let the set of deterministic unknown parameters be written as

$$ \theta = [\psi, \zeta, s_3^T \mathbf{J}^T, \ldots, s_i^T \mathbf{J}^T, \ldots, s_i^T \mathbf{J}^T, \sigma_h^T, \sigma]^T $$

(43)

and the set of random parameters as

$$ \mathbf{h} = \begin{bmatrix} \hat{h} \\ \tilde{h} \end{bmatrix} $$

(44)

Define $\mathbf{y} = [\mathbf{y}_1^T, \ldots, \mathbf{y}_L, \mathbf{y}_L^T, \ldots, \mathbf{y}_L^T]^T$ and $\mathbf{n} = [\mathbf{n}_1^T, \ldots, \tilde{\mathbf{n}}_L^T, \mathbf{h}_1^T, \ldots, \tilde{\mathbf{n}}_L^T]^T$. To alleviate notational burden in the ensuing derivations, the short notation $\mathbf{z}_i$ is used to replace $\mathbf{z}_0(\mathbf{z}_0^* + \mathbf{z}_i^*)$ in (20).

Based on the assumptions in the last paragraph of Sect. 2, the conditional probability density function (pdf) of measurements $\mathbf{y}_i$ on one realization of channel coefficients, is given by

$$ g_1(\mathbf{y}|\mathbf{h};\theta) = \frac{1}{(\pi \sigma)^N_L} \exp \left\{ -\frac{1}{\sigma} \sum_{i=1}^L \| \mathbf{y}_i - \mathbf{G}_i \mathbf{s}_i \mathbf{F} \mathbf{h} \|^2 \right\} $$

(45)

and the pdf of channel coefficients, is given by

$$ g_2(\mathbf{h}) = \frac{1}{(\pi)^L \prod_{i=0}^{L-1} \sigma_i} \exp \left\{ -\frac{1}{\sigma} \sum_{i=0}^{L-1} \frac{|h_i|^2}{\sigma_i} \right\} $$

(46)

Then the joint pdf of the noisy data $\mathbf{y}_i$ and channel coefficients, is equal to

$$ f(\mathbf{y}, \mathbf{h};\theta) = g_1(\mathbf{y}|\mathbf{h};\theta)g_2(\mathbf{h}). $$

(47)

After dropping constant terms, the negative log-likelihood function can be obtained as the following

$$ \mathcal{L} = -\log f(\mathbf{y}, \mathbf{h};\theta) $$

$$ = NL \log \sigma + \frac{1}{\sigma} \sum_{i} \| \mathbf{y}_i - \mathbf{G}_i \mathbf{s}_i \mathbf{F} \mathbf{h} \|^2 + \sum_{i=0}^{L-1} \log \sigma_i + \sum_{i=0}^{L-1} \frac{|h_i|^2}{\sigma_i}. $$

(48)

Using (48), the hybrid Fisher information matrix (HFIM) (defined in (4.609) of [18]) can be represented in the following block matrix format:
Using the symbols defined in the above paragraph, the first-order partial derivatives of subcarrier, and to the active subcarrier with index \( p \), \( \mathbf{u}_p^H \) to be the row of \( \mathbf{Q}_f \) corresponding to the same active subcarrier, and

\[
e_{x,i} = \left[ \mathbf{u}_{-p/2}^H \cdot S_i^* \left( \frac{\partial \mathbf{G}}{\partial x} \right)^H \mathbf{g}_{i,-p/2}, \ldots, \mathbf{u}_{p/2}^H \cdot S_i^* \left( \frac{\partial \mathbf{G}}{\partial x} \right)^H \mathbf{g}_{i,p/2} \right]^H
\]

where

\[
x = \psi \text{ or } \zeta, \ i = 3, 4, \ldots, I
\]

\[
e_n = [e_{x,3}^H, \ldots, e_{x,I}^H]^H
\]

\[
\mathbf{Q}_{\psi, \zeta} = \sum_{i=1}^{I} \left[ \text{Tr}[S_i^* \left( \frac{\partial \mathbf{G}}{\partial \psi} \right)^H \mathbf{G} \frac{\partial \mathbf{S}_i}{\partial \psi} \cdot \mathbf{Q}_f] \right. \left. + \text{Tr}[S_i^* \left( \frac{\partial \mathbf{G}}{\partial \zeta} \right)^H \mathbf{G} \frac{\partial \mathbf{S}_i}{\partial \zeta} \cdot \mathbf{Q}_f] \right]
\]

\[
\mathbf{Q}_3 = \text{diag} \left[ \mathbf{J} [\mathbf{G}_f^H \mathbf{G}_f] \right] \cup \mathbf{Q}_f^T \cup \ldots \cup \mathbf{J} [(\mathbf{G}_f^H \mathbf{G}_f) \cup \mathbf{Q}_f] \cup \mathbf{J} \]

\[
\mathbf{Q}_1 = (\mathbf{J} \otimes \mathbf{I}_{L-2}) \left[ e_{\psi} \ e_{\zeta} \right].
\]

Using the symbols defined in the above paragraph, the first-order partial derivatives of \( \mathcal{L} \), for deterministic parameters, evaluated at true values, can be obtained. They are presented in Appendix C. Using (22)–(25), (27) and \( \bar{z}_1 \bar{z}_2 = z_1 (z_2)^* \), the top-left block of the HFIM in (49) can be given by

\[
\mathbf{J}_{\theta \theta} = 2 \frac{1}{\sigma} \mathbf{Q}_{\psi, \zeta} \mathbf{Q}_1^H + \mathbf{J}_{\theta \psi} \mathbf{J}_{\psi \psi}^{-1} \mathbf{J}_{\theta \psi}^T
\]

where \( n_1 = (I - 2)(P - K) \).

In the matrix in (55), the elements for \( \sigma_l, \ l = 0, \ldots, L - 1 \) and \( \sigma \), are not correlated with the elements for other deterministic parameters.

Expectations of the products of first-order partial derivatives, between deterministic parameters and random parameters (i.e., channel coefficients), can also be obtained. All those products include terms either in the form \( n_1(k)h_i h_2 \) or \( n_1(k_1)n_2(k_2)h_1 \). Hence, from (26), one can prove that

\[
\mathbf{J}_{\theta h} = E \left\{ \frac{\partial \mathcal{L}}{\partial \mathbf{h}} \frac{\partial \mathcal{L}}{\partial \mathbf{\theta}^T} \right\} = 0_{2L \times n_1}, \quad \mathbf{J}_{\theta h} = E \left\{ \frac{\partial \mathcal{L}}{\partial \mathbf{h}} \frac{\partial \mathcal{L}}{\partial \mathbf{\theta}^T} \right\} = 0_{n_1 \times 2L}.
\]
Equation (56) indicates that the HCRBs in (49) for channel coefficients are not related to the HCRBs for deterministic parameters, i.e., the off-diagonal blocks of the HFIM in (49) are zero. Expectations of the products of first-order partial derivatives, between random parameters, can also be developed, leading to

$$ J_{hh} = \begin{bmatrix} \frac{2}{\sigma^2} (I_L + \sigma^2 (D_{\sigma_n})^{-1}) & 0_{L \times L} \\ 0_{L \times L} & \frac{2}{\sigma^2} (I_L + \sigma^2 (D_{\sigma_n})^{-1}) \end{bmatrix}. $$

(57)

The HCRB matrix (defined in (4.610) of [18]) is the inverse of the matrix in (49). Due to the results in (55) and (56), the HCRB matrix can be written as a matrix with two diagonal blocks:

$$ H_{\theta,h} = \begin{bmatrix} (J_{\theta\theta})^{-1} & 0_{n_1 \times 2L} \\ 0_{2L \times n_1} & (J_{hh})^{-1} \end{bmatrix}. $$

(58)

Define by $\hat{\theta}$ an estimate of the deterministic parameter set $\theta$ and $\partial z = \partial z_1, \ldots, \partial z_D$ for a random variable set $z = [z_1, \ldots, z_D]^T$. The sample space of $\tilde{y}$ is $R^{2NI}$ and that of $\tilde{h}$ is $R^{2L}$ where $R^n$ denotes the $n$-dimensional Euclidean space. Then the parameters in $\theta$ are called unbiased if

$$ E\{\hat{\theta} - \theta\} = \int_{R^{2NI}} \int_{R^{2L}} \hat{\theta} f(\tilde{y}, \tilde{h}, \theta) \partial \tilde{y} \partial \tilde{h} = 0_{n_1 \times 1}. $$

(59)

Note that $\hat{\theta}$ is a function of the noisy data symbol $\tilde{y}$ and (random) channel $\tilde{h}$, i.e., $\hat{\theta} = \hat{\theta}(\tilde{y}, \tilde{h})$. The following lemma describes the conditions for the top-left block matrix in the HCRB in (58) to be the lower bound of the estimates of deterministic parameters based on the model (20).

**Lemma 1** If (i) the regularity conditions in Assumptions 1–2 (in Section II) of [14] are satisfied and (ii) estimates $\hat{\theta}$ of deterministic parameters $\theta$ are unbiased, then the estimation error variance matrix for deterministic parameters is bounded from below by $(J_{\theta\theta})^{-1}$, i.e.,

$$ E\{(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T\} \geq (J_{\theta\theta})^{-1}. $$

(60)

**Proof** See the proof of (43) in [14]. The proof is provided in Lemma 6 in [14]. \qed

Lemma 6 in [14] is proven using some definitions and results of measure theory. Those definitions and results are beyond the scope of this paper, and hence will not be discussed. Interested readers can refer to [14] for details.

In the proof of Lemma 6 in [14], the following two assumptions are used: (i)

$$ \lim_{\tilde{h}_1 \to \pm \infty} \int_{R^{2NI}} \int_{R^{2L-1}} \frac{\hat{h}(\tilde{y}, \tilde{h}) - \hat{h}}{f(\tilde{y}, \tilde{h}, \theta | \tilde{h}_1) \partial \tilde{y} \partial \tilde{h}} = 0_{2L \times 1} $$

(61)

$$ \forall l \in [0, 1, \ldots, L - 1], $$

$$ \lim_{\tilde{h}_2 \to \pm \infty} \int_{R^{2NI}} \int_{R^{2L-1}} \frac{\hat{h}(\tilde{y}, \tilde{h}) - \hat{h}}{f(\tilde{y}, \tilde{h}, \theta | \tilde{h}_2) \partial \tilde{y} \partial \tilde{h}} = 0_{2L \times 1} $$

(62)

$$ \forall l \in [0, 1, \ldots, L - 1]; $$
where $\mathbf{h}_1^\cdot$ is the subvector of $\mathbf{h}$ without $\tilde{h}_i$ and $\mathbf{h}_2^\cdot$ is the subvector of $\mathbf{h}$ without $\tilde{h}_j$, and (ii) the unbiasedness condition for (random parameters) channel coefficients

$$E(\hat{\mathbf{h}} - \mathbf{h}) = \int_{\mathbb{R}^{2L}} \int_{\mathbb{R}^{2L}} (\hat{\mathbf{h}} - \mathbf{h}) f(\mathbf{y}, \mathbf{h}; \theta) \, d\mathbf{y} \, d\mathbf{h} = 0_{2L \times 1} \quad (63)$$

where $\hat{\mathbf{h}} = \hat{\mathbf{h}}(\mathbf{y}, \mathbf{h})$ is an estimate (vector) of $\mathbf{h}$. The conditions in (61)–(63) are required for the estimation covariance matrix for all parameters (including channel coefficients) to be bounded from below by the HCRB matrix when off-diagonal blocks of (49) are not zero in general cases. However, given the zero off-diagonal blocks of the HFIM in (49) in this paper, those conditions can be dropped.

From the joint pdf $f(\mathbf{y}, \mathbf{h}; \theta)$ in (47), one can also obtain the marginal pdf by integration with respect to channel coefficients

$$f_i(\mathbf{y}; \theta) \overset{\text{def}}{=} \int_{\mathbb{R}^{2L}} f(\mathbf{y}, \mathbf{h}; \theta) \, d\mathbf{h} \quad (64)$$

From (64), one may be able to develop another CRB. It will be called marginal CRB (MaCRB). In [14], the MaCRB was discussed and the conditions for the MaCRB to be equal to the HCRB, are given. Generally speaking, an explicit expression of the function (64) is difficult to obtain. Thus the development of the MaCRB is not a trivial task. Due to this reason, the MaCRB will neither be considered in this paper.

Theorem 4 and Corollary 5 of [14] state that if and only if

$$\frac{\partial L}{\partial \mathbf{w}} = E \left\{ \frac{\partial L}{\partial \mathbf{w}} \frac{\partial L}{\partial \mathbf{w}^T} \right\} (\hat{\mathbf{w}} - \mathbf{w}) \quad \text{a.s.} \quad (65)$$

where $\mathbf{w} = [\theta^T, \mathbf{h}^T]^T$, $\hat{\mathbf{w}}$ is an estimate vector of $\mathbf{w}$ given in a simulation run, and a.s. means almost surely, then the HCRB can be achieved by the hybrid ML/MAP estimator. In this paper, this condition is not satisfied, hence the HCRB is not achievable.

For the computational ease, from (49), (55) and (56), one can obtain the following compact HCRB matrix for the CFO and SFO only

$$\mathbf{H}_{\psi, \xi} = \frac{\sigma}{2} \left[ \sum_{i=1}^I \left[ \text{Tr}[\mathbf{Q}_{\psi, i}^H \mathbf{D}_{a_i, i}] \left[ \text{Tr}[\mathbf{Q}_{\psi, i, \xi}^H \mathbf{D}_{a_i, i}] \right] \right] \right. $$

$$- \left. \left[ \sum_{i=3}^I \left[ \mathbf{e}_{\psi, i}^H \mathbf{J}_{\psi, i}^T \right] \left[ \mathbf{J}((\mathbf{G}_i^H \mathbf{G}_i) \odot (\mathbf{Q}_f)^T) \mathbf{J}^T \right]^{-1} \left[ \mathbf{J}_{\epsilon, i} \mathbf{J}_\epsilon \right] \right] \right]^{-1} \quad (66)$$

Note that $\mathbf{G}_i^H \mathbf{G}_i$ can be simplified to $\Omega_i \Psi_i \mathbf{V} \Omega_i$.

## 5 Numerical Study

In this section, numerical evaluations will be used to compare the HCRB against the estimation accuracies of the methods in [1–3, 5, 6]. The estimation accuracy is measured by the mean squared error (MSE) of estimates yielded by a method in multiple runs of Monte-carlo simulation.

The IEEE 802.11ac standard parameters are adopted: $N_{cp} = 16$, $N = 64$, $P = 56$ [16, Table 22-5, p. 244]; and $K = 4$ pilot symbols are used and pilot subcarrier indices in (3)
are $[p_1, p_2, p_3, p_4] = [-21, -7, 7, 21]$ [16, p. 289]. Information-bearing symbols are randomly taken from the QPSK constellation for each data symbol. Pilot symbols are randomly chosen from the BPSK constellation for all data symbols. Both were fixed in all simulation runs. The modulus of QPSK symbols are taken to be equal to 1 for convenience. The CFO $\psi = 10^{-1}$ and the SFO $\zeta = 10^{-4}$ were also fixed in all simulation runs. The value of the SFO satisfies (6).

From (35), one can find that $I_0 = 7$. As pointed out in the last paragraph of Sect. 3, $s_1 = s_2$. Hence, in all simulation examples, at least eight data symbols will be used. Through numerical calculation, it is found that $s_i, i = 2, \ldots, I_0 + 1$ are independent which satisfies the second requirement of Theorem 1.

In various simulation runs, (varying) Rayleigh channels were used. Each Rayleigh channel contains $L = 12$ taps with power profiles equal to $\sigma_i = E[|h_i|^2] = e^{-i/12} / \sum_{k=0}^{11} e^{-i/12}$ for $l = 0, 1, \ldots, 11$. The signal-to-noise ratio is equal to $SNR = 10 \log_{10}[E[|x_i|^2]/(N\sigma^2)] = 10 \log_{10}[P \sum_{l=0}^{11} E[|h_i|^2]/(N\sigma^2)] = 10 \log_{10}[P/(N\sigma^2)]$ (dB) where $x_i$ is given in (12).

Due to high sensitivity of eigendecomposition to model mismatch, in the presence of noise, the MSEs of the CFO and SFO estimates of the procedures in Sect. 3.2, are far poorer than that given by the methods in [1–3, 5, 6]. If plotted in the same figures, the MSEs of the CFO and SFO estimates of the procedures in Sect. 3.2, are far poorer than that given by the methods in [1–3, 5, 6]. If plotted in the same figures, the behaviors of all the other methods will not be viewed clearly. Hence those procedures will not be considered in simulation comparison.

To show how much the HCRB can be further approached, another method is also considered. Let $\mathbf{E}$ and $\mathbf{V}$ be defined in the same way as $\mathbf{E}$ and $\mathbf{V}$ but parameterized with $\psi', \zeta'$, and $P_l^\perp = \mathbf{I} - \mathbf{E}'\mathbf{V}'((\mathbf{V}')^H\mathbf{V}')^{-1}(\mathbf{V}')^H(\mathbf{E}')^*$. Then the new method yields the CFO and SFO estimates as

$$\text{arg min}_{\psi', \zeta'} \sum_{i=1}^{l} y_i^H P_l^\perp \psi' y_i.$$  \hfill (67)

The evaluation of the function in (67) is performed for nine candidate parameter sets over a $3 \times 3$ uniform grid bounded by the true values of the CFO and SFO and the corresponding MLC estimates. This new method is called the least squares (LS) method, because the function in (67) is a concentrated least squares function of the data in (20). Clearly, the performance of the LS method will be no worse than the MLC method. The MLC estimates are chosen here because the MLC method is the most accurate among the three. One should note that the LS method can not be used in practice because the true values of the CFO and SFO are unknown.

The methods in [1–3, 5, 6], require the squared magnitudes of channel frequency responses at pilot subcarriers. In simulation, channel coefficients are determined as the average of that from the first two data symbols based on (41) with $x_i$ substituted by $y_i$.

In the $i$-th simulation run, (1) a coarse CFO estimate $\hat{\psi}'$ is obtained from the correlation between the measurements of the first two data symbols, based on the same principle as in (22) of [7]; (2) $z_{\psi, \zeta, N, \psi}' = \mathbf{E}'(\hat{\psi}')$ is multiplied to $y_i$ to obtain a compensated measurement vector as in (34) of [7] where $\mathbf{E}(\hat{\psi}')$ has the same structure as $\mathbf{E}$ but parameterized in terms of $\hat{\psi}'$; and (3) estimates of the RCFO ($= \psi - \hat{\psi}'$) and SFO are obtained by using the methods in [1–3]. Denote by $\hat{\epsilon}'$ an estimate of the RCFO given by a method, in the same simulation run as $\hat{\psi}'$. The CFO estimation error is then given by $\sqrt{\psi - (\hat{\psi}' + \hat{\epsilon}')^2}$. Thus the CFO MSE is calculated by $\sqrt{\sum_{i=1}^{N_{\text{sim}}} (\hat{\psi}' + \hat{\epsilon}' - \psi)^2/N_{\text{sim}}}$ where $N_{\text{sim}} = 2000$ is the total number of simulation runs. The SFO MSE is calculated by
\[ \sum_{i=1}^{N_{sim}} (\hat{\zeta}_i - \zeta)^2 / N_{sim} \] where \( \hat{\zeta}_i \) is an estimate of \( \zeta \) generated along with \( \hat{\epsilon}_i \) in the \( i \)-th simulation run.

The initial time instant \( I_{ini} \) is introduced in (13). In [19], the initial time instant was modeled as time delay for a single access system and treated as an unknown deterministic parameter. But in this paper, most methods considered in simulation are independent of this parameter, hence it is treated as a known deterministic parameter. Its impact on the HCRB values is first investigated.

As shown in (22) of [7], the coarse CFO estimate is approximately (for small noise) irrelevant to the initial time instant. For \( I_{ini} = 1, 2, \ldots, 9 \), the MSEs of the coarse CFO estimates are 2.34e-07, 2.33e-07, 2.32e-07, 2.33e-07, 2.35e-07, 2.37e-07, 2.38e-07, 2.36e-07, 2.34e-07, respectively, when \( \text{SNR} = 40 \text{ dB} \), and \( I = 8 \). The performances of the methods in [1–3, 5, 6], are also approximately irrelevant to \( I_{ini} \). The MSEs of estimates given by those and the LS methods along with the HCRBs are shown in Figs. 1 and 2 for \( \psi \) and \( \zeta \). From these results, one can see that the six methods have nearly constant performances while the HCRBs vary significantly (around 6 dB for \( \psi \) and \( \zeta \)). The HCRB values are also calculated for \( I_{ini} = 10, \ldots, 30 \), and it is found that the highest HCRB values occur at \( I_{ini} = 5 \). Therefore, in the remaining part of this section, \( I_{ini} = 5 \) will be used.

Figures 3 and 4 depict the HCRB along with the CFO and SFO MSEs for the SLC, MLCP, MLC, SIC, SLCP and LS methods, for \( I = 8 \). The SIC method offers better
performance that the SLC method for the CFO estimation; but the SIC method is the worst for the SFO estimation. The SLCP method offers a performance similar to the MLCP method for both CFO and SFO estimation. The MLC method has the better performance for the CFO and SFO estimation than the SLC, SLCP, MLC, MLCP methods. Its performance is further improved by the LS method. The LS CFO and SFO MSEs are approximately 11 dB and 12 dB higher than the corresponding HCRBs. Figures 5 and 6 depict the HCRB along with the MSEs for the SLC, MLCP, MLC, SIC, SLCP and LS methods, for varying $I$, at SNR = 20 dB. Similar observations can be obtained on the comparative performances of the six methods. The difference between the HCRBs and the LS MSEs slightly increases with the number of OFDM symbols.

6 Conclusions

For carrier frequency and sampling clock frequency offset estimation, the hybrid CRB has been developed.

For the identification of the carrier frequency and sampling clock frequency offsets, some necessary conditions have been found. One condition is the minimum number of data symbols required for a subspace method proposed in Sect. 3. When the number
of subcarriers is 64, and the number active subcarriers is 56, this minimum number is equal to 7. Furthermore, given two training symbols, a realization of random channel coefficients and information-bearing symbols can also be identified.

The properties of the hybrid CRB have been studied. It is found that the hybrid CRB is sensitive to the initial time instant. Hence, to assess the performances of initial-time-instant-insensitive methods, the highest hybrid CRB values for a set of initial time instant candidates should be used as a fair bound. This observation has never been reported in the literature.

The hybrid CRB has been used to assess performance limits of five existing methods. The estimation accuracies of those five methods do not reach the hybrid CRB. Further effort is required to develop new methods to improve the estimation accuracies of the CFO and SFO.

**Appendix A: Proof of Theorem 1**

Rewrite $\sum_{i=1}^{I_0} (\Omega_i S_i)\Phi\Phi^H(\Omega_i S_i)^* = Q_1^H Q_1$ where

Fig. 3 HCRB, MSEs of estimates of $\psi$, versus SNR, for the SLC, MLCP, MLC, SIC, SLCP and LS methods. $I = 8 (I_0 = 7), I_{ini} = 5$
Let \( p = \begin{bmatrix} 1, z - p (1 + \frac{N}{2}) \\ \vdots \\ z - (N - P - 1) p (1 + \frac{N}{2}) \end{bmatrix} \), then \( H = \begin{bmatrix} 1 - \frac{P}{2}, \ldots, 1 - 1, 1, \ldots, 1 - \frac{P}{2} \end{bmatrix} \). Further let \( \Omega_i = \text{diag}[\hat{s}_{i,P/2}, \ldots, \hat{s}_{i-1}, \hat{s}_{i,1}, \ldots, \hat{s}_{i,P/2}] \). Then one can write

\[
Q_1 = \begin{bmatrix}
\Phi^H(\Omega_i S_i)^* \\
\vdots \\
\Phi^H(\Omega_i S_i)^*
\end{bmatrix}.
\]

(68)

Let \( \mu_p = [1, z_0^{-p(1+\zeta)}, \ldots, z_0^{-(N-P-1)p(1+\zeta)}] \), then \( \Phi^H = [\mu_{-P/2}, \ldots, \mu_{-1}, \mu_1, \ldots, \mu_{P/2}] \). Further let \( \Omega_i S_i = \text{diag}[\hat{s}_{i,P/2}, \ldots, \hat{s}_{i-1}, \hat{s}_{i,1}, \ldots, \hat{s}_{i,P/2}] \). Then one can write

\[
Q_1 = \begin{bmatrix}
\hat{s}_{1-P/2}^* \mu_{-P/2} \\
\hat{s}_{2-P/2}^* \mu_{-P/2} \\
\vdots \\
\hat{s}_{P/2}^* \mu_{P/2} \\
\hat{s}_{1-P/2}^* \mu_{-P/2} \\
\hat{s}_{2-P/2}^* \mu_{-P/2} \\
\vdots \\
\hat{s}_{P/2}^* \mu_{P/2}
\end{bmatrix}.
\]

(69)

Note that \( Q_1 \) is an \( I_0(N - P) \times P \) matrix.

Firstly, from (35), one knows that \( I_0(N - P) \geq P \) (the number of rows of \( Q_1 \) is larger than or equal to that of columns) and \( (I_0 - 1)(N - P) < P \). Thus if (35) is not satisfied, \( Q_1 \) will have fewer rows than columns and its rank will be less than \( P \).

Secondly, if (35) is satisfied but \( [\hat{s}_1, \hat{s}_2, \ldots, \hat{s}_{I_0}] \) do not have a full column rank, at least one of its column vectors is a linear combination of others. Then one can apply

\[ \text{Springer} \]
elementary row-transformation on $Q_1$ such that a set of consecutive $N-P$ rows are equal to zero and the number of non-zero rows is less than $P$. The rank of $Q_1$ is thus less than $P$.

**Appendix B: Proof of Theorem 2**

Let $E_0$ and $V_0$ be the submatrices of $E$ and $V$, consisting of the top $P+1$ rows of the respective matrices. Define

$$\Theta = \text{diag} \begin{bmatrix} z_0^{-(P/2)(1+\zeta)} & \cdots & z_0^{-(1)(1+\zeta)} & \cdots & z_0^{(P/2)(1+\zeta)} \end{bmatrix}$$

$$\phi_k = \begin{bmatrix} z_0^{k-(P/2)(1+\zeta)} & \cdots & z_0^{k-(1)(1+\zeta)} & \cdots & z_0^{k(P/2)(1+\zeta)} \end{bmatrix}^T$$

$$\mathcal{H} = \text{diag} \begin{bmatrix} H_{-P/2}, \ldots, H_{-1}, H_1, \ldots, H_{P/2} \end{bmatrix}.$$  

Then $x_i^{(k)}$ defined in (31) can be written as
Now one can obtain that

\[
\begin{align*}
\mathbf{x}_i^{(k)} &= \frac{(n_0 + N_{\psi})((1 + \zeta)\psi)}{z_0} \mathbf{E}_0 \mathbf{V}_0 \mathbf{\Theta}^{k-1} \mathbf{\Omega} \mathbf{S}_i \mathbf{F} \mathbf{h}_0 \mathbf{z}_0^{(k-1)\psi(1+\zeta)} \\
&= \frac{(n_0 + N_{\psi})((1 + \zeta)\psi)}{z_0} \mathbf{E}_0 \mathbf{V}_0 \mathbf{\mathcal{H}} \cdot \mathbf{S}_i \mathbf{\Omega} \mathbf{\Phi}^{k-1} \mathbf{z}_0^{(k-1)\psi(1+\zeta)}. 
\end{align*}
\]  

(72)

Fig. 6 HCRB, MSEs of estimates of $\zeta$, versus the number of data symbols, for the comm and LS methods. SNR = 20 dB. $I_{ini} = 5$

Now one can obtain that

\[
\mathbf{X}_i = [\mathbf{x}_i^{(1)}, \mathbf{x}_i^{(2)}, \ldots, \mathbf{x}_i^{(N-P)}] = \frac{(n_0 + N_\psi)(1+\zeta)\psi}{z_0} \mathbf{E}_0 \mathbf{V}_0 \mathbf{\mathcal{H}} \cdot \mathbf{S}_i \mathbf{\Omega} \mathbf{\Phi} \mathbf{Q}_1
\]

(73)

where $\mathbf{\Phi} = [\mathbf{\phi}_0, \mathbf{\phi}_1, \ldots, \mathbf{\phi}_{N-P-1}]$, $\mathbf{Q}_1 = \text{diag}[1, \zeta_0^{\psi(1+\zeta)}, \ldots, \zeta_0^{(N-P-1)\psi(1+\zeta)}]$ and construct the following covariance matrix

\[
\sum_{i=1}^{I_0} \mathbf{X}_i \mathbf{X}_i^H = \mathbf{E}_0 \mathbf{V}_0 \mathbf{\mathcal{H}} \cdot \sum_{i=1}^{I_0} \mathbf{(S}_i \mathbf{\Omega}_i) \mathbf{\Phi} \mathbf{\Phi}^H \mathbf{(\Omega}_i \mathbf{S}_i)^* \cdot \mathbf{(E}_0 \mathbf{V}_0 \mathbf{\mathcal{H}})^H.
\]

(74)

In this lemma, since the necessary condition of Theorem 1 is assumed to be met, then the middle matrix in (74)

\[
\sum_{i=1}^{I_0} \mathbf{(S}_i \mathbf{\Omega}_i) \mathbf{\Phi} \mathbf{\Phi}^H \mathbf{(\Omega}_i \mathbf{S}_i)^*
\]

(75)
is of full rank $P$. Under the assumption (6), $V_0$ in (74) is guaranteed to have full column rank $P$ (so is $V$ in (20)). The matrices $E_0$ is of full column rank $P$. Furthermore, when $H_p \neq 0 \forall p \in P$, the matrix in (33) is of rank $P$. Thus $E_0 V_0 H$ is of full column rank $P$.

The theorem is therefore proved.

Appendix C: Expressions of First-Order Partial Derivatives

First-order partial derivatives evaluated at true values are equal to:

$$
\frac{\partial L}{\partial \psi} = -\frac{2}{\sigma} \sum_i h^H F^H S_i^* D_{i,\psi}^H (y_i - G_i S_i F h) = -\frac{2}{\sigma} \sum_i h^H \cdot F^H S_i^* D_{i,\psi}^H \cdot n_i
$$

\[\text{def} = -\frac{2}{\sigma} \sum_i h^H Q_{\psi, i} n_i, \quad D_{\psi, i} = \frac{\partial G_i}{\partial \psi}, \quad Q_{\psi, i} = \frac{\partial G_i}{\partial \psi} S_i F\]

(76)

$$
\frac{\partial L}{\partial \zeta} = -\frac{2}{\sigma} \sum_i h^H \cdot F^H S_i^* D_{i,\zeta}^H \cdot n_i
$$

\[\text{def} = -\frac{2}{\sigma} \sum_i h^H Q_{\zeta, i} n_i, \quad D_{\zeta, i} = \frac{\partial G_i}{\partial \zeta}, \quad Q_{\zeta, i} = \frac{\partial G_i}{\partial \zeta} S_i F\]

(77)

$$
\frac{\partial L}{\partial \mathbf{s}_i} = -\frac{2}{\sigma} \mathbf{H}^* G_i^H n_i, \quad i = 3, \ldots, I
$$

(78)

$$
\frac{\partial L}{\partial \mathbf{s}_i} = -\frac{2}{\sigma} j \mathbf{H}^* G_i^H n_i, \quad i = 3, \ldots, I
$$

(79)

$$
\frac{\partial L}{\partial \sigma} = \frac{NI}{\sigma} - \frac{1}{\sigma} \sum_i \|n_i\|^2
$$

(80)

$$
\frac{\partial L}{\partial \sigma_l} = \frac{1}{\sigma} - \frac{|h_l|^2}{(\sigma_l)^2}, \quad l = 0, 1, \ldots, L - 1
$$

(81)

$$
\frac{\partial L}{\partial \mathbf{h}} = -\frac{2}{\sigma} \sum_i \mathbf{F}^H S_i^* G_i^H \cdot n_i + 2 \mathbf{D}_{\sigma_h}^{-1} \mathbf{h}
$$

\[\text{def} = -\frac{2}{\sigma} \sum_i Q_{h, i}^H n_i + 2 \mathbf{D}_{\sigma_h}^{-1} \mathbf{h}, \quad Q_{h, i} = G_i S_i F\]

(82)

$$
\frac{\partial L}{\partial \mathbf{h}} = -\frac{2}{\sigma} \sum_i -j Q_{h, i}^H n_i + 2 -j \mathbf{D}_{\sigma_h}^{-1} \mathbf{h}
$$

(83)
**Funding** Open Access funding enabled and organized by CAUL and its Member Institutions. No funding was received for this project.

**Code Availability** The MATLAB codes used during the current study are available from the corresponding author on request.

**Data Availability** The datasets generated during the current study are available from the corresponding author on request.

**Declarations**

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

**References**

1. Tsai, P., Kang, H., & Chieuh, T. (2005). Joint weighted least-squares estimation of carrier-frequency offset and timing offset for OFDM systems over multipath fading channels. *IEEE Transactions on Vehicular Technology*, 54(1), 211–223.
2. Morelli, M., & Moretti, M. (2010). Fine carrier and sampling frequency synchronization in OFDM systems. *IEEE Transactions on Wireless Communications*, 9(4), 1514–1524.
3. Murin, Y., & Dabora, R. (2016). Low complexity estimation of carrier and sampling frequency offsets in burst-mode OFDM systems. *Wireless Communications and Mobile Computing*, 16(9), 1018–1034.
4. Kim, Y.-H., & Lee, J.-H. (2011). Joint maximum likelihood estimation of carrier and sampling frequency offsets for OFDM systems. *IEEE Transactions on Broadcasting*, 57(2), 277–283.
5. Wang, X., & Hu, B. (2014). Low-complexity estimator for carrier and sampling frequency offsets in OFDM systems. *IEEE Communications Letters*, 18(3), 503–506.
6. Zhou, X., Zheng, K., Wu, B., & Zhao, H. (2020). Successive interference cancellation-based weighted least-squares estimation of carrier and sampling frequency offsets for WLANs. *IEEE Access*. https://doi.org/10.1109/ACCESS.2020.3034582.
7. Cheng, Q. (2018). Joint estimation of carrier and sampling frequency offsets using OFDM WLAN preamble. *Wireless Personal Communication*, 98(2), 2121–2161.
8. Kumar, M., & Dabora, R. (2019). A novel sampling frequency offset estimation algorithm for OFDM systems based on cyclostationary properties. *IEEE Access*, 7, 100692–100705.
9. You, Y.-H., & Lee, K.-T. (2010). Accurate pilot-aided sampling frequency offset estimation scheme for DRM broadcasting systems. *IEEE Transactions on Broadcasting*, 56(4), 558–563.
10. You, Y.-H., Jung, Y.-A., & Paik, J.-H. (2019). Joint estimation of symbol timing and sampling frequency offset for CDD-OFDM-Based DRM systems. *IEEE Transactions on Broadcasting*, 65(2), 333–339.
11. Nguyen-Le, H., Le-Ngoc, T., & Ko, C. C. (2009). RLS-Based joint estimation and tracking of channel response, sampling, and carrier frequency offsets for OFDM. *IEEE Transactions on Broadcasting*, 55(1), 84–94.
12. Messer, H., & Adar, Y. (1989). New lower bounds on frequency estimation of a multitone random signal in noise. *Signal Processing*, 18, 413–424.
13. Miller, R. B., & Chang, C. B. (1978). A modified Cramer-Rao bound and its applications. *IEEE Transactions on Information Theory*, 24, 398–400.
14. Noam, Y., & Messer, H. (2009). Notes on the tightness of the hybrid Cramer-Rao lower bound. *IEEE Transactions on Signal Processing*, 57(6), 2074–2084.
15. Van Trees, H. L., & Bell, K. L. (2007). *Bayesian bounds for parameter estimation and nonlinear filtering/tracking*. Wiley-Inter-science; Wiley; IEEE.

16. IEEE Std 802.11ac - 2013 IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications - Amendment 4: Enhancements for Very High Throughput for Operation in Bands below 6 GHz.

17. Speth, M., Fechtel, S. A., Fock, G., & Meyr, H. (1999). Optimum receiver design for wireless broadband systems using OFDM—Part I. *IEEE Transactions on Communications*, 47(11), 1668–1677.

18. Van Trees, H. L., Bell, K. L., & Tian, Z. (2013). *Detection estimation and modulation theory: Part I, Detection, estimation, and filtering theory* (2nd ed.). Hoboken: Wiley.

19. D’Andrea, A. N., Mengali, U., & Reggiannini, R. (1994). The modified Cramer-Rao bound and its application to synchronization problems. *IEEE Transactions on Communications*, 42(2/3/4), 1391–1399.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Dr. Qi Cheng** obtained the B.E. degree from Anhui University in 1982, the M.E. degree from the Institute of Electronics of the Chinese Academy of Science in 1985, and the Ph. D. degree from the University of Melbourne in 1995. He has taught a wide range of subjects in computer, electrical and telecommunication engineering, and assumed various academic administrative roles, at the Northern Territory University and Western Sydney University. He is currently a Lecturer at the Western Sydney University. Dr. Cheng’s research interests lie in telecommunication and radar signal processing. He has published more than 60 technical papers in journals and international conferences. He is a co-editor of the book “High Resolution Signal Processing” (Editors: Y. Hua, A. Gershman, Q. Cheng) published by Marcel Dekker in 2004. He was session chairs of a number of international conferences.

**Dr. Mosin (Andre) Russell** received his master of engineering (honours) degree from University of Western Sydney in 2014 and PhD degree in computer vision from Western Sydney University in 2019. His research interests include digital image and signal processing, computer vision applications, machine learning and computer graphics. He is currently a lecturer at the Western Sydney University.