Approximate Maximum Likelihood Estimation of Integer Carrier Frequency Offset in OFDM Systems

Jungwon Lee, Hui-Ling Lou, and Dimitris Toumpakaris
Marvell Semiconductor, Inc., 700 First Avenue, Sunnyvale, CA 94089, USA
Email: {junglee, hlou, dimitris}@marvell.com

Abstract—In orthogonal frequency division multiplexing (OFDM) systems, the carrier frequency offset can be normalized by the subcarrier bandwidth and then be divided into the integer and fractional part. This paper focuses on the estimation of the integer carrier frequency offset and derives approximate maximum likelihood (ML) estimators in the presence and in the absence of pilot subcarriers. Simplified approximate ML estimators with reduced complexity are also proposed. The performance of the proposed estimators is compared with that of existing estimators, and it is shown by simulation that the proposed estimators outperform existing estimators.

I. INTRODUCTION

Recently, orthogonal frequency division multiplexing (OFDM) has been used widely for wireless communication systems in frequency selective channels [1]. By dividing a frequency selective channel into many subchannels, OFDM systems can remove inter-symbol interference (ISI) readily. Moreover, the modulation of an OFDM system can be implemented using the computationally efficient fast Fourier transform (FFT). In fact, OFDM systems are already widely used in practical wireless systems including wireless local area networks (WLAN), fixed wireless access, digital audio broadcasting (DAB), and digital video broadcasting (DVB) systems.

Similar to single carrier systems, OFDM systems are affected from the carrier frequency offset caused by the Doppler shift and oscillator instabilities. In OFDM systems, the carrier frequency offset is usually represented relative to the subcarrier bandwidth and is divided into the integer part and the fractional part. The fractional part results in signal-to-noise ratio (SNR) reduction and bit error rate (BER) degradation, whereas the integer part causes a cyclic shift of the subcarriers [2].

Various algorithms have been proposed in the literature for the estimation of the fractional carrier frequency offset [2]–[5] and the integer frequency offset [2], [6]–[8]. This paper focuses on the estimation of the integer frequency offset and derives approximate maximum likelihood (ML) estimators in the presence and in the absence of pilot subcarriers. Simplified approximate ML estimators are also developed to reduce the computational complexity.

The remainder of the paper is organized as follows. Section II describes the OFDM system model in the presence of a carrier frequency offset. Section III derives approximate ML estimators of the integer carrier frequency offset and develops simplified approximate ML estimators. In Section IV, the proposed estimators are compared with existing estimators. Section V presents simulation results, which show the superior performance of the proposed estimators compared to existing estimators. Finally, Section VI concludes the paper.

II. SYSTEM MODEL

In an OFDM system, the m-th OFDM symbol $x_m[n]$ is generated by performing the N-point inverse fast Fourier transform (IFFT) on the information symbols $X_m[k]$ for $k = 0, 1, \cdots, N - 1$ and adding $N_g$ cyclic prefix samples to the front of the IFFT of $X_m[k]$. The OFDM symbol $x_m[n]$ is transmitted through a channel $h_m[n]$ over a carrier and is corrupted by Gaussian noise $z_m[n]$. The channel $h_m[n]$ is assumed to be block stationary, i.e., time-invariant during each OFDM symbol. At the receiver, the passband signal is downconverted to baseband using a local oscillator. When the local oscillator frequency $f_l$ is not matched to the carrier frequency $f_c$ of the received signal, a carrier frequency offset $\Delta f = f_c - f_l$ will appear. Moreover, there may also be a phase offset $\Delta \theta$. The received symbol $y_m[n]$ can then be represented as

$$y_m[n] = e^{j[2\pi f (n+m(N+N_g))T+\Delta \theta]} (h_m[n] * x_m[n]) + z_m[n], \quad (1)$$

where $T$ is the sampling period, and $z_m[n]$ is a zero-mean complex-Gaussian random variable with variance $\sigma_z^2$ and is independent of the transmit signal and the channel.

The frequency offset can be represented with respect to the subcarrier bandwidth $1/NT$ by defining the normalized carrier frequency offset $\epsilon$ as

$$\epsilon \triangleq \frac{\Delta f}{1/NT} = \Delta f NT. \quad (2)$$

The normalized carrier frequency offset can be divided into an integer part $l$ and a fractional part $\hat{\epsilon}$ such that $-1/2 \leq \hat{\epsilon} < 1/2$:

$$\epsilon = l + \hat{\epsilon}, \quad (3)$$

where $l$ is an integer. The estimation of the fractional part has been investigated in [2]–[5]; many of the techniques do not require the knowledge of the integer part of the carrier frequency offset. Since the fractional part can be estimated and corrected before the estimation of the integer part, this paper focuses on the estimation of the integer carrier frequency offset in the absence of the fractional carrier frequency offset.

When the fractional frequency offset is equal to zero, the discrete Fourier transform (DFT) of $y_m[n]$ can be expressed
\[ Y_m[k] = e^{j(2\pi lm \alpha + \Delta \theta)} H_m[k - l] X_m[k - l] + Z_m[k], \]  

where \( \alpha = N_0 \), and \( H_m[k] \) and \( Z_m[k] \) are the DFTs of \( h_m[n] \) and \( z_m[n] \), respectively. In (4), it was assumed that \( H_m[k] \) and \( X_m[k] \) are periodic with period \( N \) to simplify the notation. Similarly, \( Y_m[k] \) is also assumed to be periodic throughout this paper. As can be seen in (4), the integer frequency offset \( l \) causes a cyclic shift and a phase change proportional to the OFDM symbol number.

The integer frequency offset \( l \) can be estimated either with or without the help of pilot subcarriers. To handle both pilot-aided estimation and blind estimation, the estimation problem is considered in a general context. Let \( S_p \) and \( S_d \) be the set of indices for pilot subcarriers and data subcarriers, respectively. Let the number of elements in \( S_p \) and \( S_d \) be \( N_p \) and \( N_d \), respectively. Depending on the particular OFDM system and OFDM symbol, \( N_p \) or \( N_d \) can be zero. Then, the transmit symbols satisfy the following relationship:

\[ X_{m-1}[k] X_m[k] = \begin{cases} A_m[k] & \text{for } k \in S_p, \\ B_m[k] & \text{for } k \in S_d. \end{cases} \]

where \( A_m[k] \) is a sequence known to the receiver, and \( B_m[k] \) is a random sequence unknown to the receiver. In this paper, it is assumed that the variance of \( A_m[k] \) is equal to that of \( B_m[k] \), which is set to one without loss of generality.

### III. APPROXIMATE MAXIMUM LIKELIHOOD ESTIMATORS

In this section, approximate ML estimators of the integer carrier frequency offset are derived for the additive white Gaussian noise (AWGN) channel. From (4), for the AWGN channel, the received signal can be represented as follows:

\[ Y_m[k] = e^{j(2\pi lm \alpha + \Delta \theta)} X_m[k - l] + Z_m[k]. \]

Since the phase offset \( \Delta \theta \) is unknown to the receiver, \( Y_m[k] \) is multiplied by \( Y_{m-1}[k] \) to remove \( \Delta \theta \) from the desired signal \( X_m[k - l] \) and \( X_{m-1}[k - l] \):

\[ V_m[k] \triangleq Y_{m-1}[k] Y_m[k] = e^{j2\pi lm} X_{m-1}[k - l] X_m[k - l] + Z'_m[k], \]

where the noise \( Z'_m[k] \) is

\[ Z'_m[k] = e^{-j(2\pi (m-1) \alpha + \Delta \theta)} X_{m-1}[k - l] Z_m[k] + e^{j(2\pi m \alpha + \Delta \theta)} X_m[k - l] Z'_{m-1}[k] + Z_m[k]. \]

For the derivation of a closed-form approximate ML estimator, it is assumed that \( Z'_{m-1}[k] Z_m[k] \) is negligible compared to the other terms in (8). Then \( Z'_m[k] \) will approximately follow a Gaussian distribution. This approximation becomes accurate when the SNR is high. Moreover, the variance of \( Z'_m[k] \) is approximated as \( \sigma_{Z'_m}^2 \approx 2 \sigma_Z^2 \).

In this paper, it is assumed that \( V_m[0], V_m[1], \cdots, \) and \( V_m[N - 1] \) are the observations used by the integer frequency offset estimator. Let the observation vector \( \mathbf{V}_m = [V_m[0] V_m[1] \cdots V_m[N - 1]] \). From now on, the OFDM symbol number \( m \) is omitted for the sake of notational simplicity. The ML estimate of the integer frequency offset \( l \) given the observation \( \mathbf{V} \) is the integer \( n \) that maximizes the following statistic:

\[ T_1(n) \triangleq f(\mathbf{V}|l = n), \]

where \( f \) is the conditional probability density function (pdf) of \( \mathbf{V} \) given \( l = n \). Since the observation \( \mathbf{V} \) depends not only on the integer frequency offset \( l \) but also on the values of the data symbols \( B[k] \) for \( k \in S_d \), the conditional pdf \( f(\mathbf{V}|l = n) \) can be rewritten as follows:

\[ T_1(n) = \sum_{\mathbf{b} \in S_d} f(\mathbf{V}|\mathbf{B} = \mathbf{b}, l = n) P\{\mathbf{B} = \mathbf{b}\}, \]

where \( \mathbf{B} = [B[k_1] B[k_2] \cdots B[k_{N_d}]] \) for \( k_i \in S_d \), and \( \mathbf{b} = [b[k_1] b[k_2] \cdots b[k_{N_d}]] \) represents the actual value taken by the random vector \( \mathbf{B} \). For the concreteness of the paper, it is assumed that the data subcarriers are modulated using quadrature phase shift keying (QPSK). The ML estimator for the case of other constellations can be derived in a similar way. Since \( P\{\mathbf{B} = \mathbf{b}\} = 2^{-2N_d} \) for QPSK, the following \( T_2(n) \) can be used instead of \( T_1(n) \) for the ML estimation:

\[ T_2(n) \triangleq \sum_{\mathbf{b} \in S_d} f(\mathbf{V}|\mathbf{B} = \mathbf{b}, l = n). \]

When the received signals in each subcarrier are independent of one another, the conditional pdf \( f(\mathbf{V}|\mathbf{B} = \mathbf{b}, l = n) \) becomes

\[ f(\mathbf{V}|\mathbf{B} = \mathbf{b}, l = n) = \prod_{k \in S_p} f_{Z^*}(V[k + n] - e^{j2\pi \alpha} A[k]) \cdot \prod_{k \in S_d} f_{Z}(V[k + n] - e^{j2\pi \alpha} b[k]) \]

where

\[ f_{Z^*}(x) \approx \frac{1}{\pi \sigma_{Z^*}^2} e^{-|x|^2/2\sigma_{Z^*}^2} = \frac{1}{2\pi \sigma_Z^2} e^{-|x|^2/2\sigma_Z^2} \]

By simple algebraic manipulation, it can be shown that

\[ f_{Z}(x) \approx \frac{1}{2\pi \sigma_{Z^*}^2} e^{-|x|^2/2\sigma_{Z^*}^2} \]

where \( \Re[x] \) denotes the real part of \( x \), and \( f_{Z^*}(V[k + n] - e^{j2\pi \alpha} A[k]) \) can be similarly represented. Using these expressions in (12) and removing all factors that are independent of \( n \), it can be shown that the ML estimator approximately maximizes

\[ T_3(n) \triangleq \sum_{\mathbf{b} \in S_p} \left( \prod_{k \in S_p} e^{\Re[V[k + n] - e^{j2\pi \alpha} A^*[k]]} \right) \cdot \left( \prod_{k \in S_d} e^{\Re[V[k + n] - e^{j2\pi \alpha} b^*[k]]} \right). \]
Since each $b[k_i]$ for $k_i \in S_d$ can take only the values $\pm 1$ and $\pm j$ and the $b[k_i]'$s are independent of one another, $T_3(n)$ can be rewritten as

$$T_3(n) = \prod_{k \in S_p} e^{\frac{2j\pi n \alpha k}{\sigma_k^2}} \cdot \prod_{k \in S_d} 2 \left( \cosh \left( \frac{2\pi n \alpha k}{\sigma_k^2} \right) + \cosh \left( \frac{2\pi n \alpha k}{\sigma_k^2} \right) \right) \left( \Im \{V[k+n]e^{-j2\pi n\alpha}\} \right) \right),$$

where $\Im \{x\}$ denotes the imaginary part of $x$. By taking the logarithm of $T_3(n)$, removing the constant terms, and multiplying by $\sigma_k^2$, an approximate ML pilot-aided estimator (PAE) of the integer frequency offset can be expressed as (17), which can be found at the top of the next page. In the absence of pilot subcarriers, the approximate ML blind estimator (BE) can be expressed as (18).

This approximate ML estimation makes intuitive sense: it uses the fact that the integer carrier frequency offset causes a phase shift over two OFDM symbols and a cyclic shift of the subcarriers. The cyclic shift of the subcarriers is reflected on the shift of the index of $V[k]$ in the calculation of (17) and (18), whereas the phase shift over two OFDM symbols is exploited as follows. In the first summation in (17), each term measures the component of the observation $V[k+n]$ for $k \in S_p$ in the direction of $e^{j2\pi n\alpha}A[k]$. In the absence of noise, each term should be equal to 1 when the estimate is equal to the actual integer frequency offset $l$. However, it will have a value less than 1 when the estimate $n$ is different from $l$. In the second summation in (17), each term measures the magnitude of the observation $V[k+n]$ for $k \in S_d$ in the direction of $\pm e^{j2\pi n\alpha}$ and $\pm e^{j(2\pi n\alpha+\frac{\pi}{2})}$ and takes the average with a function comprised of $\log$ and $\cosh$. In the absence of noise, each term should be equal to $\sigma_k^2 \log \left( \cosh \left( \frac{2\pi n \alpha k}{\sigma_k^2} \right) \right)$ when $(n-l)\alpha$ is an integer multiple of $\frac{\pi}{2}$. It will have a smaller value otherwise. The computation of (18) is essentially equivalent to that of the second summation in (17).

The approximate ML estimators can be simplified for high SNR by noting that

$$\cosh(x) + \cosh(y) = \frac{1}{2} (e^x + e^{-x} + e^y + e^{-y})$$

$$\approx \frac{1}{2} e^{\max\{|x|,|y|\}}$$

for $|x| \gg |y| > 0$ or $|y| \gg |x| > 0$. With this further approximation at data subcarriers, a simplified approximate ML PAE can be expressed as

$$\hat{i}_{\text{simp,PAE}} = \arg\max_n \left\{ \sum_{k \in S_p} \Re \{V[k+n]e^{-j2\pi n\alpha}A^*[k]\} \right\} \sum_{k \in S_d} \max \left\{ \left| \Re \{V[k+n]e^{-j2\pi n\alpha}\} \right|, \right\} \left| \Im \{V[k+n]e^{-j2\pi n\alpha}\} \right| \right\},$$

whereas a simplified approximate ML BE is

$$\hat{i}_{\text{simp,BE}} = \arg\max_n \left\{ \sum_{k \in S_p} \max \left\{ \Re \{V[k+n]e^{-j2\pi n\alpha}\}, \right\} \right\} \sum_{k \in S_d} \max \left\{ \left| \Re \{V[k+n]e^{-j2\pi n\alpha}\} \right|, \right\} \left| \Im \{V[k+n]e^{-j2\pi n\alpha}\} \right| \right\}. \tag{21}$$

For simplicity, in the remainder of the paper, the simplified approximate ML estimators are called as the simplified estimators.

IV. COMPARISON WITH EXISTING ESTIMATORS

Two of the most popular pilot-aided estimators (PAEs) and blind estimators (BEs) are Schmidl’s PAE [7] and BE [6], and Morelli’s PAE [8] and BE [8], respectively. Schmidl’s PAE is

$$\hat{i} = \arg\max_n \left\{ \sum_{k \in S_p} V[k+n]A^*[k] \right\}^2, \tag{22}$$

whereas Schmidl’s BE can be represented as

$$\hat{i} = \arg\max_n \left\{ \sum_{k \in S_d} \frac{\{V[k]e^{-j2\pi n\alpha}\}^4}{|V[k]|^4} \right\}. \tag{23}$$

Morelli’s PAE is

$$\hat{i} = \arg\max_n \left\{ \sum_{k \in S_p} \Re \{V[k+n]e^{-j2\pi n\alpha}A^*[k]\} \right\} \left\{ \sum_{k \in S_d} \frac{\{V[k+n]e^{-j2\pi n\alpha}\}^4}{|V[k]|^4} \right\} \sum_{k \in S_d} \frac{|Y_i[k+n]|^2}{|Y_i[k+n]|^2}, \tag{24}$$

and Morelli’s BE is

$$\hat{i} = \arg\max_n \left\{ \sum_{k \in S_d} \sum_{i = m-1} V[i+k+n] \right\}. \tag{25}$$

When pilot subcarriers are not employed, Schmidl’s BE exploits the fact that the integer frequency offset introduces a phase shift over two OFDM symbols. On the other hand, Morelli’s BE uses the facts that the integer frequency offset causes a cyclic shift of the subcarriers and that the unused subcarriers have less energy than the used subcarriers. By measuring the energy of the subcarriers, Morelli’s estimator can locate the position of the used subcarriers and thus determine the integer frequency offset. The approximate ML BE and the simplified BE take advantage of both phenomena caused by the integer frequency offset, i.e., a phase shift over two OFDM symbols and a cyclic shift of the subcarriers.

When pilot subcarriers are used, Schmidl’s PAE takes the correlation with the known pilot subcarrier values and measures the magnitude of the correlation. This is based on the fact that the random values in the data subcarriers will have small correlation with high probability. On the other hand, Morelli’s PAE takes the correlation with the known pilot subcarrier values multiplied by $e^{-j2\pi n\alpha}$ and measures the magnitude of the correlation. Essentially, Morelli’s PAE
\[ i_{\text{approx,PAE}} = \arg\max_n \left\{ \sum_{k \in S_p} \Re\{V[k+n]e^{-j2\pi\alpha}A^*[k]\} \right. \]
\[ \left. + \sum_{k \in S_d} \sigma_Z^2 \log \left( \cosh\left( \frac{\Re\{V[k+n]e^{-j2\pi\alpha}\}}{\sigma_Z^2} \right) \right) + \cosh\left( \frac{\Im\{V[k+n]e^{-j2\pi\alpha}\}}{\sigma_Z^2} \right) \} \right\} \].
\[ i_{\text{approx,BE}} = \arg\max_n \left\{ \prod_{k \in S_d} \left( \cosh\left( \frac{\Re\{V[k+n]e^{-j2\pi\alpha}\}}{\sigma_Z^2} \right) + \cosh\left( \frac{\Im\{V[k+n]e^{-j2\pi\alpha}\}}{\sigma_Z^2} \right) \right) \right\}. \]

relies on the fact that only the pilot subcarriers will have large magnitude of correlation with the known pilot subcarriers in the direction of the \(e^{j2\pi\alpha}\). Moreover, Morelli’s PAE uses the fact that the integer frequency offset causes a cyclic shift of the subcarriers, as does Morelli’s BE. For the pilot subcarriers, the approximate ML PAE and the simplified PAE perform the same operation as Morelli’s PAE. However, for the data subcarriers, the approximate ML PAE and the simplified PAE exploit both the phase shift property and the subcarrier rotation property unlike Morelli’s PAE.

**V. Simulation Results**

In this section, the performance of the proposed approximate ML and simplified estimators are compared to that of Schmidl’s estimators, and Morelli’s estimators, by Monte Carlo simulation. The following simulation parameters are chosen. The number \(N\) of data samples and the number \(N_u\) of cyclic prefix samples in one OFDM symbol are 256 and 14, respectively, resulting in cyclic prefix width ratio \(\alpha\) of \(N_u/N = 7/256\). The maximum magnitude of the integer frequency offset is chosen as 4. The number \(N_p\) of pilot subcarriers is fixed at 7, whereas the number \(N_d\) of data subcarriers is chosen to be either 249 or 241. Thus, the number \(N_u\) of used subcarriers is either 256 or 248.

Fig. 1 presents the probability of failure of the blind estimators for the AWGN channel. Schmidl’s, Morelli’s, the approximate ML (18), and the simplified (21) estimators are compared. The performance of the integer carrier frequency offset estimates is also compared for a multipath fading channel. The

![Fig. 1. Probability of failure of the blind estimators for the AWGN channel. Schmidl’s [6], Morelli’s [8], the approximate ML (18), and the simplified (21) estimators are compared.](image-url)
multipath channel used in the simulation is comprised of 15 paths, each of which is an independent Rayleigh fading channel. The 15 paths have an exponential power delay profile, and the root-mean-square delay spread of the multipath channel is 2 samples. The approximate ML and the simplified estimators were derived for the AWGN channel and are not guaranteed to be optimal for multipath fading channels. However, the proposed estimators still show better performance than the existing Schmidl’s and Morelli’s estimators in multipath fading channels.

Figs. 3 and 4 show the probability of failure of the blind estimators and pilot-aided estimators, respectively. The simulation results for the multipath fading channel are similar to the results for the AWGN channel except that Morelli’s BE performs worse than Schmidl’s BE for $N_u = 248$ and SNR larger than 6 dB. These figures indicate that the approximate ML and the simplified estimators developed for the AWGN channel perform better than existing estimators even for the multipath channels, especially for high SNR.

VI. CONCLUSION

This paper derived approximate ML estimators for the integer carrier frequency offset in OFDM systems for the cases when pilot subcarriers exist and when they do not. Simplified approximate ML estimators with less computational complexity were also developed for high SNR. Although the approximate and the simplified approximate ML estimators were derived for the AWGN channel, simulation results show that they outperform existing estimators not only for the AWGN channel, but also for multipath fading channels.

REFERENCES