ANALYSIS AND OPTIMIZATION OF AUTO-CORRELATION BASED FREQUENCY OFFSET ESTIMATION

I.M. Ngebani∗, J.M. Chuma† and S. Masupe‡

∗ Dept. of Information Science and Electronics Engineering, 38 Zheda Road, Zhejiang University, Hangzhou 310027, China E-mail: iboz55@gmail.com
† College of Engineering and Technology, Botswana International University of Science and Technology, Private Bag 14, Palapye, Botswana E-mail: chumaj@biust.ac.bw
‡ College of Engineering and Technology, Botswana International University of Science and Technology, Private Bag 14, Palapye, Botswana E-mail: masupes@biust.ac.bw

Abstract: In this letter, a general auto-correlation based frequency offset estimation (FOE) algorithm is analyzed. An approximate closed-form expression for the Mean Square Error (MSE) of the FOE is obtained, and it is proved that, given training symbols of fixed length \( N \), choosing the number of summations in the auto-correlation to be \( \frac{2N}{3} \) and the correlation distance to be \( \frac{2N}{3} \) is optimal in that it minimizes the MSE. Simulation results are provided to validate the analysis and optimization.

Key words: Auto-correlation, frequency offset estimation, optimization, performance analysis, un-biased estimator.

1. INTRODUCTION

Carrier Frequency Offset (CFO), caused by frequency deviation between a transmitter and a receiver exists in most communication systems and may result in severe performance degradation or even system failure. Therefore, estimation and compensation of frequency offset in communication systems is important in order to allow coherent demodulation of the transmitted signals. Compared to single-carrier modulation, Orthogonal Frequency Division Multiplexing (OFDM) is more sensitive to frequency offset because it introduces Inter-Carrier Interference (ICI) and destroys the orthogonality among sub-carriers [1]. To mitigate the negative impact of frequency offset, continuous efforts have been made to develop efficient Frequency Offset Estimation (FOE) algorithms.

FOE can be done in the time or frequency domain. In OFDM systems, time-domain algorithms are typically used to estimate the initial frequency offset and frequency-domain algorithms are used to track the residual frequency offset. Time-domain FOE algorithms generally rely on the auto-correlations between two specially designed training signal segments [2–5]. Further enhancements of utilizing training signals composed of multiple identical segments have been proposed in [7, 8]. [9] gives a comparative study of the Schmidl-Cox (SC) [5] and Morelli-Mengali (MM) [6] algorithms for frequency offset estimation in OFDM, along with a new least squares (LS) and a new modified SC algorithm. In [10], the author proposes a novel maximum likelihood (ML) based algorithm for estimating the timing offset and carrier frequency offset in OFDM systems under dispersive fading channels.

Although auto-correlation based FOE algorithms have been used in many practical systems, the performance analysis and optimization of the algorithms has not yet been thoroughly investigated. In this letter, a general auto-correlation based FOE algorithm is analyzed, a closed-form expression for the Mean Square Error (MSE) of the FOE is derived, and it is proved that if the training symbol length is fixed to be \( N \), to minimize the MSE, the optimal number of summations in the auto-correlation should be \( \frac{2N}{3} \) and the optimal auto-correlation distance equals \( \frac{2N}{3} \). This letter is organized as follows: Section 2 introduces a general auto-correlation based frequency offset algorithm. The main result is presented in Section 3. Section 4 presents simulation results and some discussions. Finally, conclusions are drawn in Section 5.

2. AUTO-CORRELATION BASED FREQUENCY OFFSET ESTIMATION

A quasi-static dispersive channel that contains \( L \) resolvable multi-paths can be denoted by \( \{ h_l \}_{l=0}^{L-1} \). Let \( s_n \) be the \( n \)-th transmitted training symbol with unit energy, then the \( n \)-th received symbol can be expressed as

\[
y_n = e^{j\theta_n} \sum_{l=0}^{L-1} h_l s_{n-l} + v_n, \tag{1}
\]

where \( v_n \) is the AWGN with zero mean and variance \( \sigma^2 \) and \( \theta_n \) is the rotation angle at the \( n \)-th symbol caused by
the frequency offset. In (1), it is assumed that the rotation angles for \( L \) consecutive symbols are approximately the same, this is valid if the frequency offset is not absurdly large.

Let \( \Delta f_s \) be the true frequency offset and \( T_s \) be the symbol interval, then \( \theta_s \) can be expressed as \( \theta_s = n\Delta \theta \), where \( \Delta \theta \) is the rotation angle per symbol, and is defined as

\[
\Delta \theta \triangleq 2\pi f_s T_s.
\]

Auto-correlation based FOE relies on training symbols of length \( N \) that are composed of multiple identical segments, each segment has \( M_s \) symbols. A sensible design should have \( M_s \gg L \).

The auto-correlation metric between \( y_n \) and \( y_{n+D_1} \) is

\[
Q(M_1) = \frac{1}{M_1} \sum_{n=1}^{M_1} (y_n^\dagger)(y_{n+D_1}),
\]

where \( ^\dagger \) denotes complex conjugation, \( D_1 \) is called the "auto-correlation distance", \( M_1 \) is the number of summations in the auto-correlation and is called the "complementary auto-correlation distance". Fig.1 illustrates the autocorrelation based FOE, from Fig.1 it is clear that \( D_1 = N - M_1 \).

Having obtained \( Q(M_1) \), the frequency offset can be estimated as \([2, 3]\)

\[
\hat{\Delta} f_s = \frac{\angle Q(M_1)}{2\pi D_1 T_s}.
\]

If \( \Delta f_s \) is in the range \((-\frac{1}{2D_1 T_s}, \frac{1}{D_1 T_s})\), equation (4) can provide correct estimation, otherwise there exists a \( 2\pi \) or multiples of \( 2\pi \) phase ambiguity. In this case, the correct rotated angle should be \( \angle Q(M_1) + 2\pi d \) instead of \( \angle Q(M_1) \), where \( d \) is an integer. To resolve the phase ambiguity, another auto-correlation metric with a shorter auto-correlation distance \( D_2 \) is used, i.e., calculating

\[
Q(M_2) = \frac{1}{M_2} \sum_{n=1}^{M_2} (y_n^\dagger)(y_{n+D_2}),
\]

where \( M_2 \) is the corresponding complementary auto-correlation distance. Clearly, the two auto-correlation metrics have the relation

\[
\frac{D_1}{D_2} \angle Q(M_2) \approx \angle Q(M_1) + 2\pi d,
\]

and the \( 2\pi d \) phase ambiguity can be estimated as

\[
\hat{d} = \left\lfloor \frac{D_1}{D_2} \angle Q(M_2) - \angle Q(M_1) \right\rfloor / 2\pi,
\]

where \( \lfloor \cdot \rfloor \) is the rounding operation. Then, the estimated frequency offset equals

\[
\hat{\Delta} f_s = \frac{\angle Q(M_1) + 2\pi \hat{d}}{2\pi D_1 T_s}.
\]

In the autocorrelation based FOE algorithm introduced above, the FOE precision is mainly determined by \( M_1 \) and the range of resolved frequency offset is determined by \( M_2 \). In the following, we analyze the performance of the auto-correlation based FOE algorithm, and show how to optimize the algorithm.

3. PERFORMANCE ANALYSIS AND PARAMETER OPTIMIZATION

For the auto-correlation based FOE algorithm, clearly, the larger the auto-correlation distance (i.e., \( D_1 \) or \( D_2 \) ) is, the finer the estimated frequency offset, and the better the performance. However, given a fixed training symbol length \( N \), large auto-correlation distances mean smaller complementary auto-correlation distances (i.e. \( M_1 \) or \( M_2 \)). The smaller the complementary auto-correlation, the lesser the number of samples used to calculate the auto-correlation metric and thus leading to poor performance. Therefore, given \( N \), there is an optimal auto-correlation distance where the MSE is minimized.

Since \( M_2 \) is only used to resolve the ambiguity, it is sufficient to choose \( M_2 \) to satisfy the following inequality

\[-\pi < 2\pi (N - M_2) \Delta f_s T_s < \pi.\]

In the following, we only focus on how to optimize the parameter \( M_1 \). We first derive the MSE of the estimated frequency offset with complementary auto-correlation distance \( M_1 \).

Because of the repeated segments, \( D_1 \) is a multiple of \( M_1 \),
The estimation of MSE of the estimated frequency offset is given by
\[ \Delta \hat{\theta} = \frac{\sum_{l=0}^{L-1} h_l s_{n+l} - l + v_{n} + D_l}{e^{i(n+D_l)\Delta \theta}. \]  

Let us define \( z_n \triangleq \sum_{l=0}^{L-1} h_l s_{n+l}. \) Assuming independent and unit energy training symbols \( s_n, \) we have \( E[|z_n|^2] = ||h||^2 \sum_{l=0}^{L-1} ||h||^2, \) and as \( M_1 \) gets large, we have the following approximation
\[ \frac{1}{M_1} \sum_{n=1}^{M_1} ||z_n||^2 \approx ||h||^2. \]  

Substituting (10) into (5) and using the above approximation, \( Q(M_1) \) can be expressed as
\[ Q(M_1) = \frac{1}{M_1} \sum_{n=1}^{M_1} ||z_n||^2 e^{iD_1 \Delta \theta} + \bar{v}(M_1) \approx ||h||^2 e^{iD_1 \Delta \theta} + \bar{v}(M_1), \]  

where \( \bar{v}(M_1) \) is called the “noise term” for FOE and is given by
\[ \bar{v}(M_1) = A + B + C, \]  

where \( A, B \) and \( C \) are defined as:
\[ A \triangleq \frac{1}{M_1} \sum_{n=1}^{M_1} \left( v_{n}^{\dagger} e^{i(n+D_l)\Delta \theta} \right), \]  
\[ B \triangleq \frac{1}{M_1} \sum_{n=1}^{M_1} \left( v_{n}^{\dagger} e^{-iD_1\Delta \theta} \right), \]  
\[ C \triangleq \frac{1}{M_1} \sum_{n=1}^{M_1} \left( v_{n}^{\dagger} v_{n+D_l}. \right) \]  

Using equation (12) and resolving the \( 2\pi d \) ambiguity, we obtain
\[ \angle Q(M_1) + 2\pi d = D_1 \Delta \theta + \alpha, \]  

where \( \alpha \) is the angle induced by noise term \( \bar{v}(M_1). \) Note that \( \alpha \neq \angle \bar{v}(M_1). \) instead, it is the angle between \( Q(M_1) \) and \( e^{iD_1\Delta \theta} \) (See Fig.2).

The estimation of \( \Delta f_0 \) in equation (8) can be derived as
\[ \Delta \hat{f}_0 = \alpha \frac{2\pi D_1 T_c}{ \Delta f_0 }, \]  

\( \Delta \hat{f}_0 \) is later shown to be an unbiased estimator, and the MSE of the estimated frequency offset is given by
\[ R \triangleq \frac{E(||\alpha||^2)}{4\pi^2 D_1^2 T_c^2}. \]  

To get optimal FOE performance, \( M_1 \) should be chosen to satisfy
\[ M_1^{opt} = \arg \min_{M_1} \{ R \}. \]  

The following theorem summarizes the main result of this letter, which gives \( M_1^{opt} \), and the minimum MSE.

**Theorem 1:** For a system with \( N \) training symbols for FOE, the optimal complementary auto-correlation distance that minimizes the MSE of the estimated frequency offset is
\[ M_1^{opt} = \left\lceil \frac{N}{3} \right\rceil \]  

and the corresponding minimum MSE is approximately
\[ R_{\min} \approx \frac{1}{8\pi^2 T_c^2} \left( N - \left\lceil \frac{N}{3} \right\rceil \right) \left( \frac{2}{SNR} + \frac{1}{SNR^2} \right), \]  

where \( SNR \triangleq \frac{||h||^2}{\sigma^2}. \)

**Proof:** Expanding the expectation of \( ||\bar{v}(M_1)||^2 \) in (13), we have
\[ E[||\bar{v}(M_1)||^2] = E \left[ ||A + B||^2 \right] + E \left[ (A + B)C^\dagger \right] \]  
\[ + E \left[ C(A + B)^\dagger \right] + E \left[ ||C||^2 \right]. \]

Since \( v_n \) is a complex Gaussian random variable with zero mean, we have \( E[ (A + B)C^\dagger] = 0 \) and \( E[ C(A + B)^\dagger] = 0. \) Therefore, \( E[||\bar{v}(M_1)||^2] \) can be simplified to
\[ E[||\bar{v}(M_1)||^2] = E \left[ ||A + B||^2 \right] + \frac{\sigma^2}{M_1}. \]  

**Case 1:** \( M_1 \leq \left\lceil \frac{N-1}{2} \right\rceil \)

In this case, there is no overlap between \( v_n \) and \( v_{n+D_l} \) for \( n = 1, 2, \cdots, M_1, \) so \( A \) and \( B \) are independent zero mean circular complex Gaussian random variables. Since \( \bar{v}(M_1) \) does not favor any specific direction, we have \( E[\alpha] = 0. \) This makes \( \Delta \hat{f}_0 \) given in equation (18) an unbiased estimator.

Based on the illustration in Fig.2, assuming \( M_1 \) is large, in high SNR scenarios, the angle \( \alpha \) can be approximated as
\[ \alpha \approx \frac{||\bar{v}(M_1)|| \sin \varphi}{||h||^2}, \]  

where \( \varphi \) is the angle between \( \bar{v}(M_1) \) and \( e^{iD_1\Delta \theta}. \) In this

It is important to note that the distribution of \( \alpha \) in equation(18) is unknown even though the first and second moments are known. Since the distribution is unknown, the CRLB cannot be derived for this dedicated case.
because $A_k + D_1 + B_k$ can be written as

$$A_k + D_1 + B_k = \frac{2\Re \left\{ i_{k+D_1} z_{k+D_1} e^{(k+D_1)i\theta} \right\} e^{iD_1i\theta}}{M_1}.$$ 

Regrouping the terms in $A + B$, we obtain

$$A + B = \sum_{n=1}^{D_1} (A_n + B_{n+M_1-D_1}) + \sum_{k=D_1+1}^{M_1} A_k + B_k - D_1,$$

where $w(M_1)$ is the summation of correlated terms and is along the direction of $e^{iD_1i\theta}$, so it has no contribution to the angle $\alpha$. Then, $\nu(M_1)$ can be re-written as

$$\nu(M_1) = \left( \sum_{n=1}^{D_1} (A_n + B_{n+M_1-D_1}) + C \right) + \nu(M_1).$$

Using similar arguments as in Case 1, we have $\mathbb{E}[\alpha] = 0$, which leads to an unbiased estimation of $\Delta f_1$ given by equation (18).

Based on the illustration in Fig.2, we can approximate the angle $\alpha$ as

$$\alpha \approx \frac{\mathbb{E}[\nu(M_1)] \sin \phi}{||h||^2},$$

where $\phi$ is the angle between $u(M_1)$ and $e^{iD_1i\theta}$.

Following the same procedure as in Case 1, we have

$$\mathbb{E} \left[ ||\nu(M_1)||^2 \right] \approx \frac{2(N - M_1)||h||^2\sigma^2}{M_1^2} + \frac{\sigma^4}{M_1},$$

and the corresponding MSE equals

$$R_2 \approx \left( \frac{2}{8\pi^2T_2^2(N - M_1) + \frac{1}{SNR}} \right).$$

Define $L(M_1) \triangleq \frac{1}{8\pi^2T_2^2M_1^2(N - M_1)^2SNR}$ and $D \triangleq L(M_1^{op})$,

where $M_1^{op}$ is given by equation (27). $R_1$ and $R_2$ given by equations (25) and (34), respectively can then be re-written as

$$R_1(M_1) = L(M_1) \left( \frac{2N}{M_1} - 2 + \frac{1}{SNR} \right),$$

$$R_2(M_1) = L(M_1) \left( \frac{2N}{M_1} - 2 + \frac{1}{SNR} \right).$$

We know that $L(M_1) \geq D$, and the minimum value of $R_1$ is $R_1^{\min} = D(2 + \frac{1}{SNR})$. To complete the proof we show that
and the MSE becomes
\[ R_2(\Delta f_s) = R_1(M_1^p) = D \left( 2 + \frac{1}{SNR} \right). \]

\[ L(M_1) \geq D \]  \hspace{1cm} \( (35) \)

\[ R_2(M_1) \geq D \left( \frac{2N}{M_1} - 2 + \frac{1}{SNR} \right) \]  \hspace{1cm} \( (36) \)

\[ R_2(M_1) \geq 2D \left( \frac{2N}{M_1} - 2 + \frac{1}{SNR} \right) \]  \hspace{1cm} \( (37) \)

\[ R_2(M_1) \geq D \left( \frac{2N}{M_1/2} - 2 + \frac{1}{SNR} \right) \]  \hspace{1cm} \( (38) \)

\[ R_2(M_1) \geq D(2 + 1/\Delta f_s) = R_1^{\min} \]  \hspace{1cm} \( (39) \)

Therefore, the minimum MSE in Case 1 is a global minimum.

4. SIMULATION VALIDATIONS AND DISCUSSION

To validate the analysis and optimization, we consider a communication system that has \( N = 500 \) symbols, \( \Delta f_s = 10kHz \), and \( 1/T_s = 1MHz \). To satisfy (9), we choose \( N_2 = 480 \) symbols.

The simulated and theoretical results of MSE vs. \( M_1 \) are shown in Fig.3. It can be seen that the MSE calculated from our analysis matches the simulated MSE very well, and the minimum MSE is achieved when \( M_1 = 167 = \left\lceil \frac{N}{2} \right\rceil \), as predicted by Theorem 1.

From Fig.3, it can be observed that the curve for \( SNR = 10dB \) is more symmetric than the curve for \( SNR = 0dB \) and the local minimum in the curve of \( SNR = 10dB \) is closer to the global minimum. This is because at high SNRs, the \( \frac{1}{SNR} \) term in (25) and (34) can be ignored and the MSE becomes 
\[ R \approx \frac{2}{SNR^2 T_s^2 M_1 (N-M_1)^2} \]  \hspace{1cm} for Case 1 and 
\[ R \approx \frac{2}{SNR^2 T_s^2 M_1 (N-M_1)(SNR)} \]  \hspace{1cm} for Case 2, respectively.

They are symmetric to the center \( M_1 = \left\lceil \frac{N-1}{2} \right\rceil \) and reach the same minimum when \( M_1 = \left\lceil \frac{N}{2} \right\rceil \) and \( M_1 = N - \left\lceil \frac{N}{2} \right\rceil \), respectively.

As a last comment, from the closed-form MSE formulas, we can see that, when \( N \) is fixed, the MSE of FOE is just a function of \( M_1 \) and \( SNR \), and is independent of \( \Delta f_s \).

5. CONCLUSION

In this letter, a general auto-correlation based FOE algorithm was analyzed, closed-form expressions of the MSE were derived, and it was proved that the optimal complementary auto-correlation distance equals \( \left\lceil \frac{N}{2} \right\rceil \), where \( N \) is the total number of training symbols. The results obtained in the letter can be of practical usage when designing training symbols in the implementation of auto-correlation based FOE algorithms.

REFERENCES


