

# Data-Aided Joint Estimation of Carrier Frequency Offset and Frequency-Selective Time-Varying Channel

Nathan Ricklin and James R. Zeidler  
School of Electrical and Computer Engineering  
University of California, San Diego  
La Jolla, California 92037  
Email: nricklin@ucsd.edu, zeidler@ece.ucsd.edu

**Abstract**—The issue of data-aided joint carrier frequency offset (CFO) and channel estimation is investigated in the presence of a rapidly time-varying frequency-selective channel. We derive novel joint CFO/channel estimators for two cases: known and unknown channel statistics. The CFO estimator that makes use of known channel statistics is shown to attain the Cramer-Rao bound. If channel statistics are unknown to the estimator then a parameterization of the time-varying channel is required and this is shown to degrade performance. Theoretical analysis of the derived estimators is provided along with simulation results.

## I. INTRODUCTION

Carrier frequency offset is a result of mismatch between the local oscillators of the transmitter and receiver of a communication system. The mismatch causes performance degradation depending upon the specific modulation format and receiver structure.

OFDM is particularly sensitive to carrier frequency offset because it relies upon the orthogonality of its subcarriers to separate its multiplexed data streams [1]. A slight offset in carrier frequency shifts the OFDM spectrum, causing a loss of orthogonality among the sub-carriers, which, upon applying the receiver FFT, results in so-called intercarrier interference. Hence, carrier frequency offset must be effectively estimated and corrected to achieve optimal system performance.

Communication systems frequently make use of a preamble block of known data as a training sequence before data is transmitted and detection is performed at the receiver. The training sequence can be used for channel impulse response estimation as well as CFO estimation.

Previous data-aided CFO estimation approaches derived for the time-invariant frequency-selective channel include [2], in which two consecutive identical symbols are transmitted, inducing a correlation structure in the received signal from which the CFO can be extracted, and [3], which presents a similar technique, but uses PN sequences to achieve a wider estimation range. In [4] an elegant framework is provided for joint ML estimation of CFO and time-invariant channel and it

is shown that the joint estimation decouples to first estimating the CFO, then the channel. Extensions of [4] include [5] and [6] which perform gradient descent on the joint likelihood function and a decision directed approach, respectively. The CFO estimator derived in [7] is also an extension of [4], with reduced complexity and a closed form approximation.

When the receiver and/or transmitter are mobile, a Doppler shift spreads the signal in frequency, and causes a time-variation of the channel taps. Because the CFO can be viewed as a special case of channel time-variation, it is more difficult to jointly estimate both the CFO and channel, and the previously mentioned estimators suffer a performance degradation.

Previous approaches to joint estimation of CFO and time-varying channel include [8], which presents a blind method and assumes channel time-variation statistics are known, and [9], which parameterizes the channel with an autoregressive model then uses the EM algorithm to determine the AR coefficients, from which a CFO estimate can be extracted. In [10], two consecutive OFDM preamble symbols comprised of CAZAC sequences are used, allowing a closed form CFO estimate.

In this work we propose a more direct approach to data-aided joint CFO and channel estimation in a frequency-selective time-varying channel. We use a framework very similar to [4], but that takes into account each channel tap at each point in time. We directly estimate the channel and the CFO via maximum likelihood when the channel statistics are not known, and via a Bayesian approach when they are.

When channel statistics are unknown, joint ML estimation faces the potential problem of having to estimate more parameters (all channel taps at all points in time plus CFO) than can be supported by the available data, also known as the “identifiability problem” [11]. This problem arises in the time-varying channel, in contrast to the time-invariant channel, because simply taking more data only leads to more points in time at which the channel needs to be estimated. Thus the ML estimator is found to require a channel parameterization.

This paper is organized as follows: In section II we define the system model for the time-varying frequency-

selective channel then in sections III and IV we derive joint CFO/channel estimators for the cases of unknown and known channel statistics, respectively. We derive the CRLB and high SNR approximations to the mean and variance of the estimators in section V, and finally present simulation results in VI.

## II. SYSTEM MODEL

For a sequence  $x[n]$ ,  $-L + 1 \leq n \leq M - 1$ , transmitted through a frequency-selective time-varying channel, the sampled complex-baseband received sequence  $y[n]$  is given by

$$y[n] = e^{j2\pi\theta n/M} \sum_{l=0}^{L-1} h[l, n] x[n-l] + w[n], \quad (1)$$

where  $\theta$  is the carrier frequency offset normalized to the block symbol rate (i.e. measured in OFDM subcarrier spacings),  $M$  is the block size (number of OFDM subcarriers),  $L$  is the number of channel taps,  $w[n]$  is zero-mean complex Gaussian circularly-symmetric noise and  $h[l, n]$  is the  $l$ th channel tap at time  $n$ .

The values of  $x[n]$  at times  $n < 0$  affect the received sequence at times  $n \geq 0$  due to the length of the channel, and in the case of OFDM, they would be part of the cyclic prefix. The received sequence  $y[n]$  can be stacked and represented in vector form as

$$\mathbf{y} = \Omega \sum_{l=0}^{L-1} \mathbf{A}_l \mathbf{h}_l + \mathbf{w}, \quad (2)$$

where  $\mathbf{y} = [y[0], y[1], \dots, y[M-1]]^T$ ,  $\mathbf{w} = [w[0], w[1], \dots, w[M-1]]^T$ , the CFO matrix  $\Omega$  is  $M$  by  $M$  and diagonal with entries  $[1, e^{j2\pi\theta/M}, \dots, e^{j2\pi\theta(M-1)/M}]$ ,  $\mathbf{h}_l = [h[l, 0], h[l, 1], \dots, h[l, M-1]]^T$  fully describes the time variation of the  $l$ th channel tap, and  $\mathbf{A}_l$  are diagonal matrices dependant on the transmitted pilot sequence with entries  $[x[0-l], x[1-l], \dots, x[M-1-l]]$ .

Intuitively, we can see that estimation of the carrier frequency offset is confounded by the fact that it is hard to tell it apart from the time-varying channel. Note that carrier frequency offset is in fact just a special case of channel time variation; we could consider the  $l$ th channel tap at time  $n$  to be  $e^{j2\pi\theta n/M} h[l, n]$ , or in the vector case, the  $l$ th channel tap vector could be  $\Omega \mathbf{h}_l$ .

## III. CFO ESTIMATION WITH UNKNOWN CHANNEL STATISTICS

We can further rewrite equation (2) as

$$\mathbf{y} = \Omega \mathbf{A} \mathbf{h} + \mathbf{w}, \quad (3)$$

where  $\mathbf{A}$  is the stacked  $\mathbf{A}_l$  matrices,  $\mathbf{A} = [\mathbf{A}_0, \mathbf{A}_1, \dots, \mathbf{A}_{L-1}]$ , and  $\mathbf{h}$  is the stacked channel tap time-variation vectors:  $\mathbf{h} = [\mathbf{h}_0^T, \mathbf{h}_1^T, \dots, \mathbf{h}_{L-1}^T]^T$ .

When the channel statistics are not known, both the channel and CFO are assumed deterministic and unknown. Estimation

takes the form of jointly maximizing the likelihood of the received data given the channel and CFO:

$$\{\hat{\theta}, \hat{\mathbf{h}}\} = \arg \max_{\theta, \mathbf{h}} p(\mathbf{y} | \theta, \mathbf{h}) \quad (4)$$

$$= \arg \min_{\theta, \mathbf{h}} \|\mathbf{y} - \Omega \mathbf{A} \mathbf{h}\|^2. \quad (5)$$

For any  $\theta$ , the  $\mathbf{h}$  that minimizes the cost function in (5) is

$$\hat{\mathbf{h}} = \mathbf{A}^\dagger \Omega^H \mathbf{y}, \quad (6)$$

where  $(\cdot)^\dagger$  denotes the pseudoinverse. Plugging (6) into (5) it follows that

$$\hat{\theta} = \arg \max_{\theta} g_1(\theta), \quad (7)$$

where  $g_1(\theta) = \mathbf{y}^H \Omega \mathbf{A} \mathbf{A}^\dagger \Omega^H \mathbf{y}$  is the cost function for estimating the CFO via joint ML in the time-varying channel.

But, since  $\mathbf{h}$  contains  $ML$  elements, the matrix  $\mathbf{A}$  is  $M$  by  $ML$ , and so  $\mathbf{A} \mathbf{A}^\dagger$  is identity and our cost function reduces to  $\mathbf{y}^H \mathbf{y}$ , rendering CFO estimation impossible. This is the ‘‘identifiability problem’’ as discussed in [11], where conditions are given that guarantee the estimability of the CFO in a static channel. Note that if the channel were time-invariant, we could simply increase the length of the pilot sequence and thereby increase the number of observed data points to mitigate this problem. However, in the the time-varying channel, increasing the length of the transmitted pilot sequence does not mitigate the identifiability problem, because as the pilot sequence gets longer, so do the lengths of the channel tap time-variation vectors. We are forced to try an alternate method.

Since it is impossible to jointly estimate the value of each channel tap at each time instant along with the CFO, we reduce the number of parameters that need to be estimated by parameterizing the channel. To do this, we project  $\mathbf{h}_l$  onto a set of  $I$  orthonormal basis vectors  $\mathbf{f}_1, \dots, \mathbf{f}_I$  that presumably can describe most of the time variation.

Grouping the basis vectors into a matrix  $\mathbf{F} = [\mathbf{f}_1, \dots, \mathbf{f}_I]$ , we can approximate the  $l$ th channel time variation vector as

$$\mathbf{h}_l \approx \mathbf{F} \alpha_l, \quad (8)$$

where  $\alpha_l$  is a length  $I$  vector of coefficients corresponding to each basis vector. Using this channel parameterization, we must now estimate the  $LI$  channel coefficients  $\alpha_0, \dots, \alpha_{L-1}$  and the CFO, a total of  $LI + 1$  parameters. This is possible if the received vector  $\mathbf{y}$  contains at least this many entries.

Rewriting (2) to take into account the approximation in (8), we obtain

$$\mathbf{y} = \Omega \sum_{l=0}^{L-1} \mathbf{A}_l \mathbf{F} \alpha_l + \mathbf{w}, \quad (9)$$

which can be rewritten similarly to (3) as

$$\mathbf{y} = \Omega \mathbf{A}_F \alpha + \mathbf{w}, \quad (10)$$

where  $\mathbf{A}_F = [\mathbf{A}_0\mathbf{F}, \mathbf{A}_1\mathbf{F}, \dots, \mathbf{A}_{L-1}\mathbf{F}]$ , and  $\alpha = [\alpha_0^T, \alpha_1^T, \dots, \alpha_{L-1}^T]^T$ . The conditional PDF of the data according to the model in (10) is

$$p(\mathbf{y}|\theta, \alpha) = \frac{1}{\pi^M |\mathbf{R}_w|} \exp [(\mathbf{y} - \Omega \mathbf{A}_F \alpha)^H \mathbf{R}_w^{-1} (\mathbf{y} - \Omega \mathbf{A}_F \alpha)] \quad (11)$$

Now, jointly maximizing this approximate likelihood of both CFO and channel parameters can be done as

$$\{\hat{\theta}, \hat{\alpha}\} = \arg \max_{\theta, \alpha} p(\mathbf{y}|\theta, \alpha) \quad (12)$$

$$= \arg \min_{\theta, \alpha} \|\mathbf{y} - \Omega \mathbf{A}_F \alpha\|^2 \quad (13)$$

and, paralleling the unparameterized case, the  $\alpha$  that minimizes (13) for any  $\theta$  is

$$\hat{\alpha} = (\mathbf{A}_F^H \mathbf{A}_F)^{-1} \mathbf{A}_F^H \Omega^H \mathbf{y}. \quad (14)$$

Note that if we choose a small enough number of basis functions,  $\mathbf{A}_F$  is a tall matrix and  $(\mathbf{A}_F^H \mathbf{A}_F)^{-1}$  exists. The CFO estimate then becomes

$$\hat{\theta} = \arg \max_{\theta} g_2(\theta), \quad (15)$$

where  $g_2(\theta) = \mathbf{y}^H \Omega \mathbf{A}_F (\mathbf{A}_F^H \mathbf{A}_F)^{-1} \mathbf{A}_F^H \Omega^H \mathbf{y}$  is the cost function for estimating the CFO via joint ML after parameterizing the time-varying channel. Here we see that the CFO estimator decouples from the channel estimator, but it depends on the particular channel parameterization that we have chosen.

Note that if the channel is assumed constant, there is only one basis vector and  $\mathbf{F}$  is simply the (scaled) all ones vector. Under this scenario  $\mathbf{A}_F$  becomes an  $M$  by  $L$  static-channel convolution matrix and the above cost function in (15) reduces to that of Morelli and Mengali's MLE#1 estimator from [4] which is optimal (achieves the CRLB) for time-invariant channels.

#### IV. CFO ESTIMATION WITH KNOWN CHANNEL STATISTICS

When the channel statistics are known, i.e. we have available to us the channel correlation matrix  $\mathbf{R}_h = E[\mathbf{h}\mathbf{h}^H]$  and the noise variance  $\sigma^2$  with  $E[\mathbf{w}\mathbf{w}^H] = \sigma^2 \mathbf{I}$ , we can apply Bayesian methods to estimate the CFO and channel. Here the channel is assumed to be a zero mean complex Gaussian circularly symmetric random variable in accordance with Rayleigh fading, but the CFO is still considered deterministic.

Assuming Rayleigh fading and uncorrelated channel taps,  $\mathbf{R}_h$  is an  $ML$  by  $ML$  block diagonal matrix

$$\mathbf{R}_h = \begin{bmatrix} \mathbf{R}_h^{(0)} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_h^{(1)} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{R}_h^{(L-1)} \end{bmatrix} \quad (16)$$

where each block on the diagonal is equal to the time correlation of a particular channel tap  $\mathbf{R}_h^{(l)} = E[\mathbf{h}_l \mathbf{h}_l^H]$ . If we assume

that each channel tap varies at the same rate, then each block in (16) is a scaled version of the same time-correlation matrix:

$$\mathbf{R}_h^{(l)} = p_l \mathcal{R}_h \quad (17)$$

where  $p_l$  is the power in the  $l$ th channel tap and the  $\{i, j\}$ th entry of  $\mathcal{R}_h$  is  $J_0(2\pi f_d T(i-j)/M)$ , according to the Jakes model.  $f_d$  is the Doppler frequency and  $T$  is the duration of the length- $M$  received data vector.

Maximum a posteriori (MAP) estimation yields

$$\{\hat{\theta}, \hat{\mathbf{h}}\} = \arg \max_{\theta, \mathbf{h}} p(\mathbf{y}|\theta, \mathbf{h}) p(\mathbf{h}) \quad (18)$$

$$= \arg \min_{\theta, \mathbf{h}} \frac{1}{\sigma^2} \|\mathbf{y} - \Omega \mathbf{A} \mathbf{h}\|^2 + \mathbf{h}^H \mathbf{R}_h^{-1} \mathbf{h}. \quad (19)$$

For any  $\theta$ , the  $\mathbf{h}$  that minimizes the cost function in (19) is

$$\hat{\mathbf{h}} = (\mathbf{A}^H \mathbf{A} + \sigma^2 \mathbf{R}_h^{-1})^{-1} \mathbf{A}^H \Omega^H \mathbf{y}, \quad (20)$$

which is also the MMSE estimate of  $\mathbf{h}$ . Plugging (20) into (19) we arrive at

$$\hat{\theta} = \arg \max_{\theta} g_3(\theta), \quad (21)$$

where  $g_3(\theta) = \mathbf{y}^H \Omega \mathbf{A} (\mathbf{A}^H \mathbf{A} + \sigma^2 \mathbf{R}_h^{-1})^{-1} \mathbf{A}^H \Omega^H \mathbf{y}$  is the cost function for estimating the CFO when the time-varying channel statistics are known.

Notes:

- There is no identifiability issue due to regularization of the inverse in  $g_3(\theta)$ . Hence, the estimator with known channel statistics does not require a channel parameterization.
- Maximizing  $g_3(\theta)$  is equivalent to minimizing  $\mathbf{y}^H \mathbf{R}_{yy}^{-1} \mathbf{y}$  which maximizes the likelihood  $p(\mathbf{y}|\theta)$ . So this Bayesian approach decouples to first finding  $\hat{\theta}$  via ML and then finding  $\hat{\mathbf{h}}$  via MMSE or MAP.

As is shown analytically and in simulations below, performance of the estimator without statistics is highly dependent upon the basis functions and the number of parameters chosen to represent each channel tap, degrading performance as compared to the estimator with known statistics.

#### V. CFO ESTIMATOR PERFORMANCE

##### A. Cramer-Rao Bound

The Cramer-Rao Lower Bound is the minimum variance attainable by any unbiased estimator for a given parameter and observations related by  $p(\mathbf{y}|\theta)$ . For a complex Gaussian data vector  $\mathbf{y} \sim CN(\mathbf{0}, \mathbf{R}_{yy}(\theta))$  with scalar parameter  $\theta$ , the CRLB is given by [12]

$$\text{Var}(\hat{\theta}) \geq \frac{1}{\text{tr} \left[ \left( \mathbf{R}_{yy}^{-1}(\theta) \frac{\partial}{\partial \theta} \mathbf{R}_{yy}(\theta) \right)^2 \right]}. \quad (22)$$

The correlation of the data is

$$\mathbf{R}_{yy}(\theta) = E[\mathbf{y}\mathbf{y}^H] = \Omega \tilde{\mathbf{R}}_{yy} \Omega^H \quad (23)$$

where  $\tilde{\mathbf{R}}_{yy} = \sum_{l=0}^{L-1} p_l \mathbf{A}_l \mathcal{R}_h \mathbf{A}_l^H + \sigma^2 \mathbf{I}$ . Taking the derivative with respect to  $\theta$  we have

$$\frac{\partial}{\partial \theta} \Omega \tilde{\mathbf{R}}_{yy} \Omega^H = \Omega \left( \mathbf{D} \tilde{\mathbf{R}}_{yy} + \tilde{\mathbf{R}}_{yy} \mathbf{D}^H \right) \Omega^H \quad (24)$$

where  $\mathbf{D}$  is a diagonal matrix with elements  $[0, j2\pi/M, \dots, j2\pi(M-1)/M]$ . Using the results of (24) and the fact that  $\mathbf{R}_{yy}^{-1}(\theta) = \Omega \tilde{\mathbf{R}}_{yy}^{-1} \Omega^H$ , we arrive at

$$\text{Var}(\hat{\theta}) \geq \frac{1}{2\text{tr}(\mathbf{D}^2 + \tilde{\mathbf{R}}_{yy}^{-1} \mathbf{D} \tilde{\mathbf{R}}_{yy} \mathbf{D}^H)}. \quad (25)$$

This bound holds for any unbiased estimator of  $\theta$  that makes use of the data  $\mathbf{y}$  to form the estimate. For example, the bound holds even if the estimator does not take advantage of, or even know, the channel statistics and correspondingly the PDF  $p(\mathbf{y}|\theta)$ . As such, it is a lower bound on the variance for all the estimators presented in this paper.

### B. Known Channel Statistics

With the same technique used in [4] and originally developed in [13] and [14], when the SNR is high we can approximate the mean and variance of the CFO estimator as

$$E[\hat{\theta}] \approx \theta - \frac{E[\frac{\partial}{\partial \theta} g(\theta)]}{E[\frac{\partial^2}{\partial \theta^2} g(\theta)]}, \quad \text{Var}(\hat{\theta}) \approx \frac{E[(\frac{\partial}{\partial \theta} g(\theta))^2]}{E[\frac{\partial^2}{\partial \theta^2} g(\theta)]^2}, \quad (26)$$

where  $g(\theta)$  is the cost function optimized by the estimator. Inserting the known-channel-statistics cost function  $g_3(\theta)$  into the mean in (26), the numerator is easily found to be zero, indicating that the estimator is unbiased.

When the cost function  $g_3(\theta)$ , which is equivalently  $\log p(\mathbf{y}|\theta)$ , is inserted into the variance in (26), it is easily manipulated into the definition of the CRLB for a single scalar parameter. Hence, the Bayesian joint estimator achieves the CFO CRLB, as expected, since it decouples to MAP or MMSE channel estimation and ML CFO estimation, which has the property of asymptotic efficiency.

### C. Unknown Channel Statistics

To find the bias of the estimator that doesn't know the channel statistics we insert  $g_2(\theta)$  into the mean in (26) to find

$$E[\hat{\theta}] - \theta = \frac{-\text{tr}[(\mathbf{D}\mathbf{Q} + \mathbf{Q}\mathbf{D}^H) \mathbf{R}'_{yy}]}{\text{tr}[(\mathbf{D}^2\mathbf{Q} + \mathbf{Q}\mathbf{D}^2 + 2\mathbf{D}\mathbf{Q}\mathbf{D}^H) \mathbf{R}'_{yy}]}, \quad (27)$$

where  $\mathbf{Q} = \mathbf{A}_F(\mathbf{A}_F^H \mathbf{A}_F)^{-1} \mathbf{A}_F^H$  and  $\mathbf{R}'_{yy} = \sum_{l=0}^{L-1} p_l \mathbf{A}_l \mathbf{R}_h \mathbf{A}_l^H$ . Equation (27) is not in general equal to zero, but it can be shown that if the pilots are purely real or imaginary, or if the channel is well parameterized, meaning that the modelling error inherent in the approximation (8) is zero, the estimator is unbiased.

To find the variance of the estimator we insert  $g_2(\theta)$  into the variance in (26), and after some manipulation, we find

$$\text{Var}(\hat{\theta}) = \frac{a + \sigma^2 b + \sigma^4 c}{d}, \quad (28)$$

where

$$a = \text{tr}^2[(\mathbf{D}\mathbf{Q} + \mathbf{Q}\mathbf{D}^H) \mathbf{R}'_{yy}] + \text{tr}[(\mathbf{D}\mathbf{Q} + \mathbf{Q}\mathbf{D}^H) \mathbf{R}'_{yy}]^2,$$

$$b = 2\text{tr}[(\mathbf{D}\mathbf{Q} + \mathbf{Q}\mathbf{D}^H)^2 \mathbf{R}'_{yy}], \quad c = \text{tr}[\mathbf{D}\mathbf{Q} + \mathbf{Q}\mathbf{D}^H]^2, \quad \text{and} \\ d = \text{tr}^2[(\mathbf{D}^2\mathbf{Q} + \mathbf{Q}\mathbf{D}^2 + 2\mathbf{D}\mathbf{Q}\mathbf{D}^H) \mathbf{R}'_{yy}]. \quad \text{In (28) we see}$$

a term that is constant in the noise variance, leading to an error floor at high SNR with a value equal to  $a/d$ . If the approximation in (8) is perfect (i.e.  $\mathbf{h}_l = \mathbf{F}\alpha_l$ ), it can be shown that  $a = 0$ , meaning there is no error floor. Thus, we conclude that the error floor is due to the modelling error in the channel parameterization.

## VI. SIMULATION RESULTS

Here we present simulations of the derived estimators in a frequency-selective time-varying channel and we compare to Morelli and Mengali's MLE#1 from [4] and to Wei, Xu, Cai, and Xu's (WXCX) estimator from [10] (with parameters  $r_1 = 2$  and  $r_2 = 20$ ). The WXCX estimator was derived for a fast time-varying channel and actually requires two separate training sequences. In our simulations we therefore give it an advantage and allow it a training sequence twice as long as that used for our estimators and the MLE#1.

For the estimator without channel statistics, we parameterize the channel with a varying number of orthogonal basis functions: from one to three. We choose these basis functions to be constant, linear, and quadratic in time. Note that the "optimal" basis functions (see [15]) cannot be chosen since they depend on the channel statistics, which are not known to the estimator. Additionally, what is optimal for channel estimation may not be optimal for CFO estimation.

We fix the length of the transmitted pilot sequence to be the sum of  $M = 256$  and the channel impulse response length  $L = 4$ . The channel taps exhibit Rayleigh fading and vary in time according to the Jakes model as in section IV, with an exponential decay power delay profile given by  $p_l = c10^{-\beta l}$  where  $c$  is a normalizing constant and  $\beta = 1/20$ .

Figures 1-3 show simulations of CFO MSE vs. SNR for normalized Doppler from  $f_d T = 0.01$  to  $f_d T = 0.3$  ( $T$  refers to the duration of an entire block of length  $M$ ). The lines show the CRLB and variance approximations while the points show simulated performance of the estimators. For all levels of Doppler, the CRLB is shown to underbound all estimators and it is achieved by the estimator with known channel statistics, as predicted in section V-B. At higher Doppler, in figures 2 and 3, the error floors of the estimators become apparent, with a higher order model used for the parameterization corresponding to a lower error floor. However, the estimators based on higher order models experience poor performance at low SNR and low Doppler due to overparameterization, i.e., the channel parameters corresponding to the higher order basis functions contain very little energy, and at low SNR estimation error is more significant than modelling error.

When the channel statistics are unknown it is difficult to choose the proper model to use for channel parameterization. Using more parameters ensures a lower error floor, but results in poor performance at low SNR. Using fewer parameters yields better performance at low SNR, but results in a higher error floor.

The WXCX estimator, which was designed for a fast varying channel, performs the same as our unknown statics estimator using a constant basis function, even though it is

able to use twice the amount of training data. This is due to the fact that in its derivation, the channel is assumed constant over each of its two pilot sequences. The extra amount of training data does not improve its performance because the second sequence merely resolves ambiguity in the CFO estimate that arises due to the CAZAC structure of the first sequence.

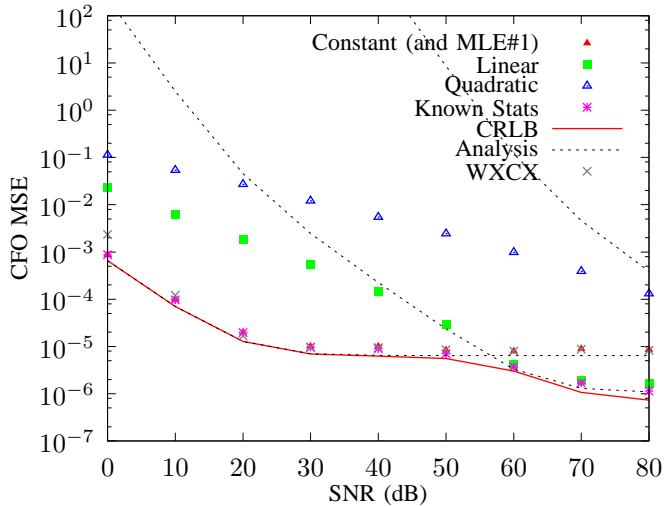


Fig. 1. CFO Estimator performance vs. SNR for  $f_d T = 0.01$

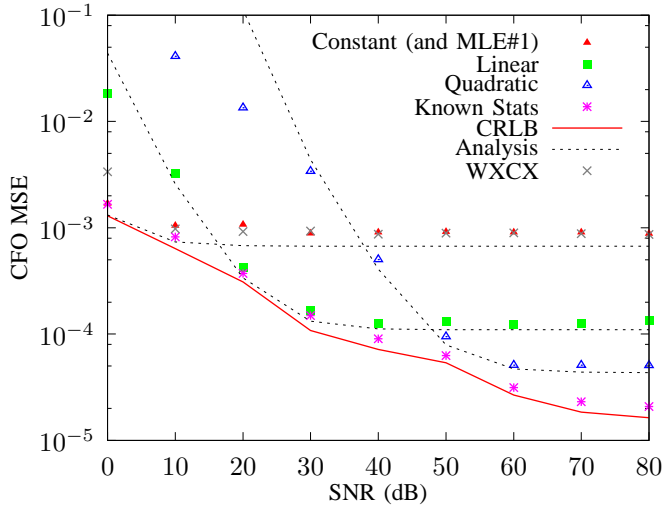


Fig. 2. CFO Estimator performance vs. SNR for  $f_d T = 0.1$

## VII. CONCLUSION

We have derived joint CFO/channel estimators for use in a time-varying frequency-selective channel for both known and unknown channel statistics. When the statistics are known, the joint estimation reduces to sequentially estimating the CFO via ML and the channel via MAP or MMSE, and the CFO estimator achieves the CRLB. When the statistics are unknown, the channel must be parameterized and performance suffers due to over or under-parameterization. Thus knowledge of the channel statistics is extremely important to achieve the best performance in all scenarios.

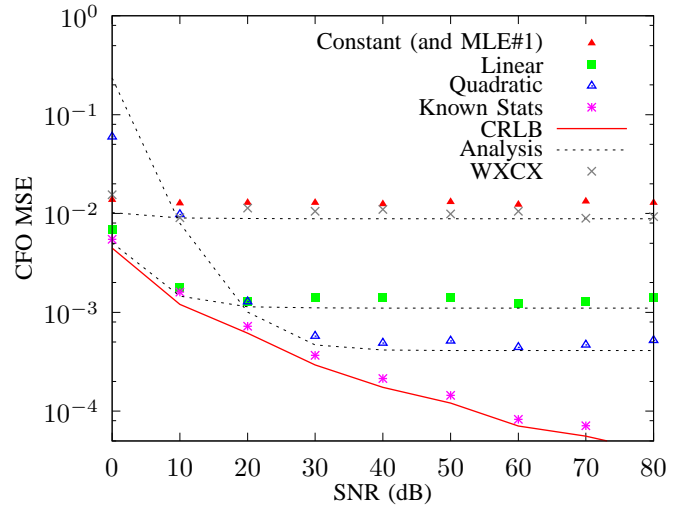


Fig. 3. CFO Estimator performance vs. SNR for  $f_d T = 0.3$

## REFERENCES

- [1] T. Pollet, M. Van Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Trans. Commun.*, vol. 43, no. 234, pp. 191–193, 1995.
- [2] P.H. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Trans. Commun.*, vol. 42, no. 10, pp. 2908–2914, 1994.
- [3] T.M. Schmidl and D.C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Trans. Commun.*, vol. 45, no. 12, pp. 1613–1621, 1997.
- [4] M. Morelli and U. Mengali, "Carrier-frequency estimation for transmissions over selective channels," *IEEE Trans. Commun.*, vol. 48, no. 9, pp. 1580–1589, 2000.
- [5] Xiaoliang Ma, H. Kobayashi, and S.C. Schwartz, "Joint frequency offset and channel estimation for OFDM," in *Proc. Globecom '03*, 1–5 Dec. 2003, pp. 15–19 Vol.1.
- [6] Tao Cui and C. Tellambura, "Robust joint frequency offset and channel estimation for OFDM systems," in *Proc. 60th IEEE Veh. Technol. Conf.*, 2004, pp. 603–607 Vol. 1.
- [7] J.H. Yu and Y.T. Su, "Pilot-assisted maximum-likelihood frequency-offset estimation for OFDM systems," *IEEE Trans. Commun.*, vol. 52, no. 11, pp. 1997–2008, 2004.
- [8] T. Lv, Hua Li, and Jie Chen, "Joint estimation of symbol timing and carrier frequency offset of OFDM signals over fast time-varying multipath channels," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4526–4535, 2005.
- [9] M. Nissila and S. Pasupathy, "Joint estimation of carrier frequency offset and statistical parameters of the multipath fading channel," *IEEE Trans. Commun.*, vol. 54, no. 6, pp. 1038–1048, 2006.
- [10] L. Wei, Y.Y. Xu, Y.M. Cai, and X. Xu, "Robust frequency offset estimator for OFDM over fast varying multipath channel," *Electronics Letters*, vol. 43, no. 6, pp. 53–54, 2007.
- [11] F. Gao and A. Nallanathan, "Identifiability of Data-Aided Carrier-Frequency Offset Estimation Over Frequency Selective Channels," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3653–3657, 2006.
- [12] Steven M. Kay, *Fundamentals of statistical signal processing: estimation theory*, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1993.
- [13] L. E. Franks, "Acquisition of carrier and timing data-I," in *NATO Advanced Study Inst. New Directions Signal Processing Commun. Contr.* Groningen, The Netherlands: Noordhoff, 1975, pp. 429–447.
- [14] M. Meyers and L. Franks, "Joint Carrier Phase and Symbol Timing Recovery for PAM Systems," *IEEE Trans. Commun.*, vol. 28, no. 8, pp. 1121–1129, 1980.
- [15] K.A.D. Teo and S. Ohno, "Optimal MMSE finite parameter model for doubly-selective channels," in *Proc. IEEE Globecom '05*, 28 Nov.-2 Dec. 2005, vol. 6.