Signal Processing ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at SciVerse ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Data-aided CFO estimators based on the averaged cyclic autocorrelation ☆

Gustavo J. González [a,*], Fernando H. Gregorio [a], Juan Cousseau [a], Stefan Werner [b], Risto Wichman [b]

[a] CONICET-Department of Electrical and Computer Engineering, Universidad Nacional del Sur, Av. Alem 1253, Bahía Blanca 8000, Argentina
[b] Aalto University, School of Electrical Engineering, P.O. Box 13000, 00076 Aalto, Finland

ARTICLE INFO

Article history:
Received 15 July 2011
Received in revised form
10 July 2012
Accepted 28 July 2012

Keywords:
CFO estimators
Cyclostationarity
Data-aided
OFDM

ABSTRACT

Wireless communication systems typically employ a repetitive preamble in each slot which is used for parameter acquisition. The repetitive preamble is useful for estimating the carrier frequency offset (CFO), usually based on the autocorrelation of the received signal. In this paper, we derive a family of novel data-aided CFO estimators. The proposed estimators are based on a new autocorrelation function which is defined using cyclostationary properties of the repetitive preamble. In contrast to previous approaches, the new estimators make use of high-order noise terms leading to an improved performance. We present a detailed analysis of the proposed estimators and provide closed-form expressions for the variance of the estimators. The new estimators are shown to outperform the existing estimators obtaining a moderate improvement at high signal to noise ratio (SNR) and a considerable improvement at low SNR, by means of a reasonable increase in computational complexity.

1. Introduction

Carrier frequency offset (CFO) estimation is an important task in wireless communication systems. This is particularly true for orthogonal frequency-division multiplexing (OFDM) because CFO destroys the orthogonality of the subcarriers [1]. CFO estimators can be generally classified into data-aided, and non-data-aided (or blind estimators). Data-aided estimators make use of a preamble, or a training sequence, for time and/or frequency synchronization [2–4], while non-data-aided estimators exploit statistical information of the signal. Cyclostationarity-based estimators have been widely studied assuming different propagation environments and modulations [5–8]. Although non-data-aided estimators are bandwidth efficient, they also require a large amount of data and consequently longer estimation time than data-aided techniques [9,10]. Therefore, advanced wireless communication systems for frequency selective channels generally employ data-aided synchronization [9–11], and the development of data-aided estimators remains an important topic.

In the acquisition stage, channel parameters are unknown and must be estimated together with time and frequency offset. Use of repetitive (or periodic) preambles can greatly simplify synchronization tasks such as carrier frequency offset and channel estimation [3]. Therefore, OFDM based wireless communication systems, e.g., WiMAX and WiFi [10], typically include in each data frame a repetitive preamble of period J (i.e., a sequence which is repeated J times).

☆ This work was partially funded by the Universidad Nacional del Sur, PGI 24/K044 and PGI24/K043, the ANPCyT, PICT-2008-0182 and PICT-2008-00104, the CONICET, PIP 112-200801-01024, the Center of Excellence in Smart Radios and Wireless Research (SMARAD), the Academy of Finland, and Finnish Funding Agency for Technology and Innovation.

* Corresponding author. Tel.: +54 291 4595101x3316; fax: +54 291 4595154.
E-mail addresses: ggonzalez@uns.edu.ar (G.J. González), Fernando.gregorio@uns.edu.ar (F.H. Gregorio), jcousseau@uns.edu.ar (J. Cousseau), stefan.werner@aalto.fi (S. Werner), risto.wichman@aalto.fi (R. Wichman).

0165-1684/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.sigpro.2012.07.032

A maximum likelihood estimator (MLE) for CFO and channel parameters has been derived in [3], in which the CFO and channel estimators are decoupled, i.e., the CFO value is computed first and then used during channel estimation. Despite the fact that the MLE is statistically efficient, its computational complexity makes it impractical. Therefore, suboptimal CFO estimators have been developed which are based on the $(J-1)$-lag autocorrelation of the training sequence. In [12], Morelli and Mengali proposed a CFO estimator that combines $J/2$ phase differences of the autocorrelation function, using the best linear unbiased estimator (BLUE). In [13], Minn proposed a two-step procedure that employs the $J-1$ autocorrelation lags together with BLUE. Although the estimator in [13] is able to exploit all available autocorrelation lags, its performance is rather poor for low signal to noise ratio (SNR) values.

This paper presents a new family of CFO estimators based on the cyclostationary properties of the periodic training sequence. We introduce a generalization of the classical autocorrelation, referred to as the *averaged cyclic autocorrelation* (ACA), which exploits the repetitive structure of the training sequence. We first reformulate the maximum likelihood estimator as a function of the ACA, and then derive three suboptimal algorithms that feature reduced computational complexity. First, exploiting the relation between the autocorrelation and the ACA, we introduce a generalization of [12] that employs the $J-1$ autocorrelation lags (instead of only $J/2$ lags). Second, using the degrees of freedom in the ACA, we derive two more algorithms that attain an accurate CFO estimate at the cost of a small increase in computational complexity. Theoretical expressions of the variance are derived for each technique, and finally, approximative versions of the proposed algorithms are derived that avoid explicit matrix inverses, which is the main source of complexity in the estimators.

The presentation is organized as follows. Section 2 introduces the signal model and reviews the algorithms [12,13]. The maximum likelihood CFO estimator based on the ACA is introduced in Section 3, and the derivation of suboptimal estimators is presented in Section 4. In this section, we also provide a detailed study of the statistics of the proposed estimators for high and low SNR regions. Section 5 presents a detailed computational complexity analysis, and the simplified versions of the suboptimal estimators. Simulation results and performance comparisons with the previous estimators are presented in Section 6, and the paper is concluded in Section 7.

2. Signal model and previous approaches

We consider a wireless communication system with a frame structure, where each frame has a periodic training sequence $x(n)$ followed by data symbols (single or multicarrier [14]). The periodic training sequence $\{x(n), -N_{cp} \le n \le N-1\}$ has period M and length $N+N_{cp}$, where the first N_{cp} samples of $x(n)$ contain the cyclic prefix. The channel $h(l)$, of length L, is assumed time invariant during the training sequence. After passing the signal through the channel, and assuming that $N_{cp} \ge L-1$, the cyclic prefix is removed and the received signal of N samples becomes

$$y(n) = e^{j(2\pi\epsilon/N)n} q(n) + w(n) \quad \text{for } n = 0, \ldots, N-1, \tag{1}$$

where $q(n) = \sum_{l=0}^{L-1} h(l)x(n-l)$ is also an M-periodic sequence of length N with $J=N/M$ periods, $w(n)$ is the white Gaussian noise with variance σ^2 and ϵ is the CFO normalized to the inter-carrier spacing. It is useful to define $s(p) = q(p)$, for $0 \le p \le M-1$, i.e., one period of $q(n)$. Assuming sufficiently long cyclic prefix to accommodate the timing offset and the channel delay spread, estimations of time offset and CFO can be decoupled and the received signal is periodic [15].

2.1. Previous approaches

CFO estimation is a nonlinear parameter estimation problem, and low-complexity CFO estimators typically follow three basic steps: an indirect measure of the CFO (usually the phase of the autocorrelation); computation of phase differences to reduce range ambiguity; and finally, a linear combination of these phase differences, usually performed using the best linear unbiased estimator (BLUE).

Following these steps, Morelli [12] and Minn [13] algorithms are based on the *sample autocorrelation* defined as

$$\hat{r}(k) = \frac{1}{N-kM} \sum_{n=kM}^{N-1} y(n)y^*(n-kM), \quad 1 \le k \le J-1, \tag{2}$$

where J is the number of periods of the training sequence. Substituting (1) into (2) results in

$$\hat{r}(k) = e^{j2\pi\epsilon k/J} \chi(k), \quad 1 \le k \le J-1, \tag{3}$$

where $\chi(k)$ is a function of the periodic training sequence $q(n)$ and $w(n)$. It is easy to note that the CFO information is divided into the $J-1$ components of $\hat{r}(k)$.

The estimator proposed by Morelli [12] can be summarized as follows:

- *Step*1: The phase of the autocorrelation function is obtained as

$$\theta(k) = \arg\{\hat{r}(k)\} = \frac{2\pi\epsilon k}{J} + \arg\{\chi(k)\}, \quad 1 \le k \le J-1. \tag{4}$$

  The estimate of ϵ from (4) is unambiguous when $|\epsilon| < J/(2k)$. Then, the estimation range varies with k, and it is not possible to combine directly the autocorrelation phases without reducing the range.
- *Step*2: To avoid the range reduction, [12] employs phase differences of $\theta(k)$, which results in

$$\theta_d(k) = [\theta(k) - \theta(k-1)]_{2\pi} = \frac{2\pi\epsilon}{J} + \gamma(k), \quad 1 \le k \le A, \tag{5}$$

  where $[\cdot]_{2\pi}$ denotes modulo 2π operation, $1 \le A \le J-1$, $\gamma(k)$ is a function that depends on $q(n)$ and $w(n)$, and $\theta(0) = 0$. As can be noted from (5), the estimation range is $|\epsilon| < J/2$.
- *Step*3: Finally, grouping the autocorrelation phase differences in the vector $\boldsymbol{\theta}_d = [\theta_d(1), \ldots, \theta_d(A)]^T$ and assuming high SNR, the BLUE of the CFO ϵ is

$$\hat{\epsilon}_M = \frac{J}{2\pi} \frac{\mathbf{1}^T \mathbf{C}_\theta^{-1} \boldsymbol{\theta}_d}{\mathbf{1}^T \mathbf{C}_\theta^{-1} \mathbf{1}}, \tag{6}$$

Please cite this article as: G.J. González, et al., Data-aided CFO estimators based on the averaged cyclic autocorrelation, Signal Processing (2012), http://dx.doi.org/10.1016/j.sigpro.2012.07.032

where $\mathbf{C}_\theta$ is the covariance matrix of $\boldsymbol{\theta}_d$. Since the derivation of the covariance matrix disregards high-order noise terms, the matrix becomes singular for $k>J/2$ and it is not possible to employ the information of autocorrelation lags larger than $J/2$. The best performance is obtained for $A=J/2$ [3].

Instead of (5), the Minn algorithm in [13] is based on the function:

$$\zeta(k)=\frac{J}{2\pi k}\theta(k)=\epsilon+\frac{J}{2\pi k}\arg\{\chi(k)\},\quad 1\le k\le J-1. \tag{7}$$

To extend the range of the estimate to $J/2$, $\zeta(1)$ is first used as a coarse CFO estimate. Then, the received signal is compensated using $\zeta(1)$ to obtain $\tilde{y}(n)=\exp(-j2\pi\zeta(1)n/N)y(n)$. Replacing $y(n)$ in (2) by $\tilde{y}(n)$, the angles $\tilde{\zeta}(k)$ can be obtained from (7), corresponding to the residual CFO, i.e., $\zeta(k)=\zeta(1)+\tilde{\zeta}(k)$ for $2\le k\le J-1$. Finally, assuming high SNR, the BLUE combines the values of $\zeta(k)$ to estimate the CFO. Three covariance matrices for $\zeta(k)$ are presented in [13], corresponding to different approaches. Two of them include high-order noise terms allowing the use of all available information ($J-1$) phases, and the third one discards the terms leading to a singular covariance matrix for $k>J/2$, as in [12]. Since the coarse CFO estimate depends only on the first autocorrelation coefficient, the technique derived in [13] is not robust at low SNR.

At this point, we make the following key observations:

- High-order noise terms in the phase statistics of the algorithm [13] allows the use of $J-1$ autocorrelation lags. As discussed in the following sections, this improves the accuracy of the CFO estimate.
- The assumptions in [12,13] are only valid for high SNR values. At low SNRs, these estimates become biased. In Appendix D we show that the BLUE based on phase differences (5) is biased for low SNR and/or high CFO.

We shall in the next section reformulate the MLE for the CFO in terms of an averaged cyclic autocorrelation (ACA). Thereafter, we derive suboptimal low-complexity CFO estimators. The proposed estimators use similar steps as [12,13], but with the important difference that they exploit the ACA instead of the traditional autocorrelation function in (2). We will see that the new estimators offer different trade-offs between accuracy and complexity.

3. Averaged cyclic autocorrelation based ML CFO estimator

The received signal in (1) can be expressed in vector form as

$$\mathbf{y}=\mathbf{D}(\epsilon)\mathbf{P}\mathbf{s}+\mathbf{w}, \tag{8}$$

where

$$\mathbf{y}=[y(0)\dots y(N-1)]^{\mathrm{T}},\mathbf{D}(\epsilon)=\mathrm{diag}[1\ e^{j2\pi\epsilon/N}\dots e^{j2\pi\epsilon(N-1)/N}],$$
$$\mathbf{P}=[\underbrace{\mathbf{I}_M\ \mathbf{I}_M\dots\mathbf{I}_M}_{J}]^{\mathrm{T}},\ \mathbf{s}=[s(0)\dots s(M-1)]^{\mathrm{T}},$$
$$\mathbf{w}=[w(0)\dots w(N-1)]^{\mathrm{T}},$$

and $\mathbf{I}_M$ is the $M\times M$ identity matrix and $s(p)$, $0\le p\le M-1$, is defined in Section 2. Matrix $\mathbf{D}(\epsilon)$ takes into account the CFO effect, whereas $\mathbf{P}$ extends $\mathbf{s}$ periodically. With the notation in (8), the joint ML estimate of $\mathbf{s}$ and ϵ is obtained by following a derivation similar to [3,16]

$$(\hat{\mathbf{s}},\hat{\epsilon}_{\mathrm{ML}})=\arg\min_{\tilde{\mathbf{s}},\tilde{\epsilon}}\|\mathbf{y}-\mathbf{D}(\tilde{\epsilon})\mathbf{P}\tilde{\mathbf{s}}\|^2, \tag{9}$$

where $\tilde{\mathbf{s}}$ and $\tilde{\epsilon}_{\mathrm{ML}}$ are the trial values of $\mathbf{s}$ and ϵ_{ML}, respectively. The estimate of $\mathbf{s}$, for a given $\tilde{\epsilon}$, is given by

$$\hat{\mathbf{s}}=\frac{1}{J}\mathbf{P}^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}(\tilde{\epsilon})\mathbf{y}. \tag{10}$$

If we then replace $\tilde{\mathbf{s}}$ with $\hat{\mathbf{s}}$ in (9), we obtain the following one-dimensional optimization problem

$$\hat{\epsilon}_{\mathrm{ML}}=\arg\max_{\tilde{\epsilon}}\{\mathrm{tr}[\mathbf{P}^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}(\tilde{\epsilon})\mathbf{y}\mathbf{y}^{\mathrm{H}}\mathbf{D}(\tilde{\epsilon})\mathbf{P}]\} \tag{11}$$

because the trace operator is invariant to circular shifts of its argument. By defining

$$\hat{r}_c(p,k)=\frac{1}{J-k}\sum_{n=0}^{J-k-1}y(nM+p)y^*((n+k)M+p), \tag{12}$$

and after some matrix manipulations, the minimization problem (11) can be stated as

$$\hat{\epsilon}_{\mathrm{ML}}=\underset{\tilde{\epsilon}}{\mathrm{argmax}}\left\{\sum_{k=1}^{J-1}(J-k)\sum_{p=0}^{M-1}\Re\{\hat{r}_c(p,k)e^{j(2\pi\tilde{\epsilon}/J)k}\}\right\}. \tag{13}$$

Eq. (13) introduces the ML estimator which is based on the cyclic autocorrelation function $\hat{r}_c(p,k)$ defined in (12), and referred to as the *averaged cyclic autocorrelation* (ACA). We see that the autocorrelation in (2) relates to the ACA in (12) as

$$\hat{r}(k)=\frac{1}{M}\sum_{p=0}^{M-1}\hat{r}_c(p,k). \tag{14}$$

Since the autocorrelation function $r(n,k)=E\{y(n)y^*(n+kM)\}$ is periodic in n, the received signal $y(n)$ in (1) is cyclostationary. As a consequence, the ACA function can be interpreted as the estimate of one period of $r(n,k)$. From (13), we see that $\hat{\epsilon}_{\mathrm{ML}}$ is periodic with period J, and that the obtained estimate is ambiguous unless ϵ is restricted in the interval defined by $|\epsilon|<J/2$. Comparing with the classic autocorrelation function in (2), which depends on a single variable, the ACA is a function of two variables (p and k).

Solving for $\tilde{\epsilon}$ in (13) requires an exhaustive grid search or the use of a DFT with huge zero padding making it impractical from the implementation point of view [3]. Therefore, suboptimal estimators trading off performance and complexity have been considered [4,12]. The new ACA function is also useful when developing low complexity CFO estimators, as discussed in the next section.

4. New averaged cyclic autocorrelation based CFO estimation algorithms

The proposed estimators are based on the ACA, and therefore we first study its mean and covariance matrix with p fixed and $1\le k\le J-1$, and with k fixed and $0\le p\le M-1$. Substituting (1) into (12) the ACA can be

written as

$$\begin{aligned}\hat{r}_c(p,k) &= \frac{1}{J-k}\sum_{u=0}^{J-k-1} e^{-j2\pi\epsilon kM/N}|s(p)|^2 \\ &\quad + e^{j2\pi\epsilon(p+uM)/N}s(p)w^*(p+(u+k)M) \\ &\quad + e^{-j2\pi\epsilon(p+(u+k)M)/N}s^*(p)w(p+uM) \\ &\quad + w(p+uM)w^*(p+(u+k)M) \\ &= e^{-j2\pi\epsilon k/J}|s(p)|^2 + w_c(p,k), \end{aligned} \tag{15}$$

where $w_c(p,k)$ is a zero-mean correlated noise process for $0\le p\le M-1$ and $1\le k\le J-1$. If required, we see from (15) that $w_c(p,k)$ can be approximated by a joint Gaussian process for the case of large $J-k$. This fact is exploited in Appendix D for studying the bias of CFO estimators based on phase differences. However, the estimators considered here only require knowledge of the first- and second-order statistics. Since $w_c(p,k)$ is not conjugate symmetric, real and imaginary parts need to be considered separately. The complete study of the first- and second-order statistics is presented in Appendix A, which is summarized below:

- For $1\le k\le J-1$ and fixed p, $\hat{r}_c(p,k)$ has mean $\boldsymbol{\mu}_k(p)$ defined in (A.5) and covariance matrix $\mathbf{C}_k(p)$ defined in (A.7).
- For $0\le p\le M-1$ and fixed k, $\hat{r}_c(p,k)$ has mean $\boldsymbol{\mu}_p(k)$ defined in (A.13) and covariance matrix $\mathbf{C}_p(k)$ defined in (A.14).

We see from (15) that the CFO can be estimated from the ACA as long as $|\epsilon|<J/(2k)$. Then, by combining p and k it is possible to extend the estimator range to $J/2$, similar to the ACA based MLE. We note that high-order noise terms are included in the ACA statistics. Furthermore, p can be considered as additional degrees of freedom offered by the ACA with respect to the classic autocorrelation function in (2). On the other hand, the combination over k can be used with any of the existing techniques to obtain full range.

In the following we present three low-complexity CFO estimators based on the ACA. These estimators can be classified into two groups depending on the variable used: *sum-based estimators* (SBEs) that use the phase of the sum of ACAs of different periods; and *direct combining estimators* (DCEs) that combine the phases of different periods. The resulting estimators can be described by three simple steps. Furthermore, we derive the theoretical variance for each algorithm, which is then compared with the Cramer–Rao lower bound (CRLB).

4.1. Sum-based CFO estimator (SBE)

The first low complexity estimator consists of the following steps:

- *Step*1: From (15), the combination over p is

$$\xi(k) = \arg\left\{\sum_{p=0}^{M-1}\hat{r}_c(p,k)\right\}, \tag{16}$$

where $\xi(k)$ is the partial CFO estimate for k.

- *Step*2: It follows from the relationship between the classic autocorrelation and the ACA of (14) that (16) is equivalent to (4). To remove the dependence on k, we take the difference between the phases at two consecutive time instants, and define

$$\boldsymbol{\xi}_d = [\xi_d(1),\ldots,\xi_d(J-1)]^T, \tag{17}$$

where

$$\xi_d(k) = [\xi(k)-\xi(k-1)]_{2\pi} \tag{18}$$

for $1\le k\le J-1$, where $\xi(0)=0$.

- *Step*3: Using the affine property in the ACA covariance matrix defined in Appendix A, the BLUE of CFO ϵ_s is

$$\hat{\epsilon}_s = -\frac{J}{2\pi}\frac{\mathbf{1}^T\boldsymbol{\Xi}_d^{-1}\boldsymbol{\xi}_d}{\mathbf{1}^T\boldsymbol{\Xi}_d^{-1}\mathbf{1}}, \tag{19}$$

$\boldsymbol{\Xi}_d$ is the covariance matrix of the random vector $\boldsymbol{\xi}_d$, see (B.9) of Appendix B, and $\mathbf{1}=[1\cdots1]^T$ $(J-1)\times 1$.

The variance of the estimate is given by

$$\mathrm{var}(\hat{\epsilon}_s) = \left(\frac{J}{2\pi}\right)^2\frac{1}{\mathbf{1}^T\boldsymbol{\Xi}_d^{-1}\mathbf{1}}. \tag{20}$$

Eqs. (19) and (20) depend on the quantity $S=\sum_{p=0}^{M-1}|s(p)|^2$ (see Appendix B). When the average energy of each symbol $x(n)$ is normalized to one, we have the mean value $E\{S\}=GM$, where $G=\sum_{l=0}^{L-1}E\{|h(l)|^2\}$. Then, as the channel is unknown, S can be replaced by the estimate $\hat{S}=\sum_{p=0}^{M-1}|\hat{r}_c(p,1)|$.

Remarks.

- The third step is the main difference between SBE and the estimator in [12]. Because of the high-order terms in $\boldsymbol{\Xi}_d$, the weighting factors in (19) employ $J-1$ autocorrelation lags, while only $J/2$ lags are used in [12].
- As usual with low-complexity CFO estimators, high SNR region is assumed for derivation of $\boldsymbol{\Xi}_d$ (as discussed in Appendix B). The behavior at low SNR is treated in Appendix D.
- Since $\xi(k)$ is equal to $\theta(k)$, $|\hat{\epsilon}_s|\le J/2$, as defined in Step 2 of Section 2.1.
- Since the ACA function is the extension of the classic autocorrelation function, the SBE can be seen as a generalization of the estimator in [12].
- The inverse of $\boldsymbol{\Xi}_d$ in (19) is the main source of complexity when compared to the classical estimation algorithms. In Section 5.1, we introduce a simplified SBE with lower complexity.

4.2. Direct combining CFO estimators (DCE)

To introduce the other two remaining estimators we first define $\alpha(p,k)=\arg\{\hat{r}_c(p,k)\}$. As stated in Appendix C, after some manipulation and assuming high SNR, we obtain

$$\alpha(p,k) = \arg\{e^{-j2\pi\epsilon k/J}|s(p)|^2 + w_c(p,k)\}$$

Please cite this article as: G.J. González, et al., Data-aided CFO estimators based on the averaged cyclic autocorrelation, Signal Processing (2012), http://dx.doi.org/10.1016/j.sigpro.2012.07.032

$$= \frac{-2\pi\epsilon k}{J} + \arctan\left\{\frac{\Im\{w_c(p,k)e^{j2\pi\epsilon k/J}\}}{|s(p)|^2 + \Re\{w_c(p,k)e^{j2\pi\epsilon k/J}\}}\right\}$$

$$\approx \frac{-2\pi\epsilon k}{J} + \frac{\Im\{w_c(p,k)e^{j2\pi\epsilon k/J}\}}{|s(p)|^2 + \Re\{w_c(p,k)e^{j2\pi\epsilon k/J}\}}$$

$$\approx \frac{-2\pi\epsilon k}{J} + \frac{\Im\{w_c(p,k)e^{j2\pi\epsilon k/J}\}}{|s(p)|^2}$$

$$\approx \frac{-2\pi\epsilon k}{J} + w_\alpha(p,k), \tag{21}$$

where $w_\alpha(p,k)$ is a correlated and real noise process for $0 \le p \le M-1$ and $1 \le k \le J-1$, of zero mean. If required, $\hat{r}_c(p,k)$ can be approximated by a joint Gaussian process for $J-k$ large, as presented in Appendix A. In that case, $w_\alpha(p,k)$ can be approximated by a joint real Gaussian process, as is discussed in Appendix C. Again, we emphasize that estimators considered here only require knowledge of first- and second-order statistics.

The second-order statistics of $\alpha(p,k)$, evaluated in Appendix C, are summarized as follows:

- For $0 \le p \le M-1$ with fixed k, $\alpha(p,k)$ has covariance matrix $\mathbf{\Theta}_p(k)$ defined in (C.3).
- For $1 \le k \le J-1$ with fixed p, $\alpha(p,k)$ has covariance matrix $\mathbf{\Theta}_k(p)$ defined in (C.6).

As shown below, these results are useful for deriving a new class of CFO estimators. Similar to the statistics used with the SBE estimator, covariance matrices $\mathbf{\Theta}_p(k)$ and $\mathbf{\Theta}_k(p)$ include high-order noise terms.

From (21) we can infer that the CFO estimate $\hat{\epsilon}$ can be obtained as long as $|\epsilon| < J/(2k)$ and, therefore, a technique to avoid range reduction must be employed. The linear dependence of ϵ in (21) suggests the application of the BLUE for combining the information contained in p and k to find the CFO estimate. Furthermore, since $\alpha(p,k)$ is approximately Gaussian for high SNR, the BLUE approximates well the MLE. In addition, we analyze the behavior for the low SNR case in Appendix D.

Depending on the order of the combination of the indexes in $\alpha(p,k)$ it is possible to obtain two different algorithms, as shown below.

B.1 Direct combining estimator Case A (DCE-A). *Step*1: From the covariance matrix $\mathbf{\Theta}_p(k)$ ((C.3) of Appendix C), we note that $\alpha(p,k)$ are statistically independent for different p. Therefore, weighting factors for combination over p can be easily obtained from the components of that covariance matrix, $[\mathbf{\Theta}_p(k)]_{p_1,p_2}$ ((C.4) of Appendix C), following [17]. The resulting weighting factors are given by

$$f(p,k) = \frac{|s(p)|^2}{(J-k)(1+\sigma^2/(2|s(p)|^2)) - (J-2k)},$$

$$g(p,k) = \frac{|s(p)|^2}{(J-k)(1+\sigma^2/(2|s(p)|^2))},$$

$$K(k) = \begin{cases} \sum_{p=0}^{M-1} f(p,k) & \text{if } k < J/2, \\ \sum_{p=0}^{M-1} g(p,k) & \text{if } k \ge J/2, \end{cases}$$

$$a(p,k) = \frac{1}{K(k)} \begin{cases} f(p,k) & \text{if } k < J/2, \\ g(p,k) & \text{if } k \ge J/2, \end{cases}$$

$$\mathbf{a}_p(k) = [a(0,k), \ldots, a(M-1,k)]^T. \tag{22}$$

Then, arranging the phases $\alpha(p,k)$ into the vector

$$\boldsymbol{\alpha}(k) = [\alpha(0,k), \ldots, \alpha(M-1,k)]^T, \tag{23}$$

we can obtain the partial estimates given by

$$\lambda(k) = \mathbf{a}_p^T(k)\boldsymbol{\alpha}(k). \tag{24}$$

*Step*2: To relax the dependence of k in the CFO estimate, we define the vector

$$\boldsymbol{\lambda}_d = [\lambda_d(1), \ldots, \lambda_d(J-1)]^T, \tag{25}$$

where

$$\lambda_d(k) = [\lambda(k) - \lambda(k-1)]_{2\pi}, \tag{26}$$

with $\lambda(0) = 0$.

*Step*3: The statistics necessary to find the CFO estimate can be obtained by noting that $a(p,k)$ defines a linear transform over $\alpha(p,k)$ for each k, and $\alpha(p,k)$ is independent for different p. Therefore, the covariance matrix of $\lambda(k)$ can be calculated using the affine transform and the independence over p, resulting in the following expression:

$$\mathbf{\Lambda}_k = \sum_{p=0}^{M-1} \mathbf{A}(p)\mathbf{\Theta}_k(p)\mathbf{A}^T(p), \tag{27}$$

where $\mathbf{A}(p) = \mathrm{diag}\{a(p,k)\}$ is a linear transform for $1 \le k \le J-1$. It is easy to show that the elements of $\mathbf{\Lambda}_k$ are

$$[\mathbf{\Lambda}_k]_{k_1,k_2} = \sum_{p=0}^{M-1} a(p,k_1)a(p,k_2)[\mathbf{\Theta}_k(p)]_{k_1,k_2}, \tag{28}$$

where $[\mathbf{\Theta}_k(p)]_{k_1,k_2}$ is defined in (C.7) of Appendix C. Considering the difference matrix transformation

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 1 \end{bmatrix} \tag{29}$$

the covariance matrix of $\boldsymbol{\lambda}_d$, defined $\mathbf{\Lambda}_d$, is given by

$$\mathbf{\Lambda}_d = \mathbf{T}\mathbf{\Lambda}_k\mathbf{T}^T. \tag{30}$$

Finally, the BLUE for CFO ϵ_a is given by

$$\hat{\epsilon}_a = -\frac{J}{2\pi}\frac{\mathbf{1}^T\mathbf{\Lambda}_d^{-1}\mathbf{\Lambda}_d}{\mathbf{1}^T\mathbf{\Lambda}_d^{-1}\mathbf{1}}. \tag{31}$$

The variance of the corresponding estimate results

$$\mathrm{var}(\hat{\epsilon}_a) = \left(\frac{J}{2\pi}\right)^2 \frac{1}{\mathbf{1}^T\mathbf{\Lambda}_d^{-1}\mathbf{1}}. \tag{32}$$

Please cite this article as: G.J. González, et al., Data-aided CFO estimators based on the averaged cyclic autocorrelation, Signal Processing (2012), http://dx.doi.org/10.1016/j.sigpro.2012.07.032

The computation of $\mathbf{\Lambda}_d$ in (30) includes $|s(p)|^2$, which is not available since the channel is unknown. For implementation purposes, $|s(p)|^2$ for a given p can be replaced by $1/(J-1)\sum_{k=1}^{J-1}|\hat{r}_c(p,k)|$.

Remarks.

- Eq. (24) employs a different combination for p than the SBE in (16) and the algorithm derived in [12]. This is facilitated by the ACA that provides an additional degree of freedom (the index p). Like in the SBE algorithm, the DCE-A algorithm also employs $J-1$ lags.
- The angles obtained from $\alpha(p,k)=\arg\{\hat{r}_c(p,k)\}$ lie between $\pm\pi$, resulting in a phase discontinuity in π. We see from (21) that for some ϵ and k, the elements of $\alpha(p,k)$ for $0\le p\le M-1$ are close to the phase discontinuity. Due to the noise, some of these values fall on different sides of the discontinuity, which decreases the performance for large CFO. If the CFO is not large, errors due to discontinuities are unlikely and the estimator has good performance as shown in Section 6.

B.2. Direct combining estimator Case B (DCE-B). *Step*1: Let us define the following phase differences for a fixed p:

$$\psi_d(p,k)=[\alpha(p,k)-\alpha(p,k-1)]_{2\pi}, \tag{33}$$

where $\alpha(p,0)=0$.

*Step*2: The statistics used to calculate the weighting factors is obtained from the covariance matrix $\mathbf{\Theta}_k(p)$ ((C.6) of Appendix C), and from the difference matrix transformation of (29), as

$$\mathbf{\Psi}(p)=\mathbf{T}\mathbf{\Theta}_k(p)\mathbf{T}^T. \tag{34}$$

Then, the weighting factors to combine $\psi_d(p,k)$ differences are given by

$$\mathbf{b}_k(p)=-\frac{J}{2\pi}\frac{\mathbf{\Psi}^{-1}(p)\mathbf{1}}{\mathbf{1}^T\mathbf{\Psi}^{-1}(p)\mathbf{1}}. \tag{35}$$

By defining $\psi_d(p)=[\psi_d(p,1),\ldots,\psi_d(p,J-1)]^T$, the partial estimates for fixed p are

$$\psi_p(p)=\mathbf{b}_k^T(p)\boldsymbol{\psi}_d(p). \tag{36}$$

*Step*3: The components of $\psi_p(p)$ are independent for different p (since $\alpha(p,k)$ are independent), giving rise to the weighting factors [17]

$$b_p(p)=\frac{\sigma_\psi^2(p)}{\sum_{p=0}^{M-1}\sigma_\psi^2(p)}, \tag{37}$$

where $\sigma_\psi^2(p)=1/(\mathbf{1}^T\mathbf{\Psi}_p^{-1}\mathbf{1})$. By grouping $\psi_p(p)$ in the vector $\boldsymbol{\psi}_p=[\psi_p(0),\ldots,\psi_p(M-1)]^T$ and $b_p(p)$ in $\mathbf{b}_p=[b_p(0),\ldots,b_p(M-1)]^T$, the final CFO estimate is given by

$$\hat{\epsilon}_b=\mathbf{b}_p^T\boldsymbol{\psi}_p. \tag{38}$$

The variance is also simplified by the independence condition and given by

$$\mathrm{var}(\hat{\epsilon}_b)=\frac{1}{\sum_{p=0}^{M-1}\sigma_\psi^2(p)}. \tag{39}$$

As with the DCE-A algorithm, $|s(p)|^2$ is replaced in a practical implementation, by the average of the ACA over k.

Remarks.

- When the CFO is small, the modulo-2π operator of (33) can be ignored.
- Combination over k is not as robust as in DCE-A or SBE, because it depends on a single component of p rather than a weighted sum. This can be seen by comparing (C.7) (Appendix C) with (B.7) (Appendix B) and (28). This becomes critical when CFO is large.
- The analysis of the bias for low SNR derived in Appendix D applies also to DCE algorithms.

5. Implementation issues

The algorithms derived in Section 4 require the inverse of the phase-difference covariance matrix to combine the information over k. In the following, we modify the algorithms to avoid the matrix inverse and discuss the computational complexity.

5.1. Approximative algorithms

The complexity of the algorithms can be reduced if the BLUE weights are replaced by the approximate weights given by

$$w_{ap}(k)=\frac{-J}{2\pi}\frac{(J-k)^2}{\sum_{k=1}^{J-1}(J-k)^2}=\frac{-1}{2\pi}\frac{6(J-k)^2}{(J-1)(2J-1)}. \tag{40}$$

The approximative weighting factors are based on the assumption that the phase differences in (18), (26) and (33) are independent. From [17], it is known that if the samples of a linear statistical model are independent, the weights are proportional to the normalized energy of each sample. Since $J-k$ in (12) is the normalization coefficient for the ACA, the selection of $w_{ap}(k)$ is natural.

Grouping the weights in the vector $\mathbf{w}_{ap}=[w_{ap}(1)\ldots w_{ap}(J-1)]^T$, we propose three approximative estimators by modifying (19), (31) and (36) as

$$\hat{\epsilon}_s^{ap}=\mathbf{w}_{ap}^T\boldsymbol{\xi}_d \tag{41}$$

$$\hat{\epsilon}_a^{ap}=\mathbf{w}_{ap}^T\boldsymbol{\lambda}_d \tag{42}$$

$$\psi_p^{ap}(p)=\mathbf{b}_k^T(p)\boldsymbol{\psi}(p) \tag{43}$$

respectively. In addition, to obtain the CFO estimate using the DCE-B, it is necessary to replace (38) by $\hat{\epsilon}_b^{ap}=\mathbf{c}^T\boldsymbol{\psi}_p^{ap}$, where $\boldsymbol{\psi}_p^{ap}=[\psi_p^{ap}(0),\ldots,\psi_p^{ap}(M-1)]^T$, $c(p)=|s(p)|^2/\sum_{p=0}^{M-1}|s(p)|^2$, and $\mathbf{c}=[c(0),\ldots,c(M-1)]^T$. In this case it is assumed that $\boldsymbol{\psi}_p^{ap}$ are independent for different p. The approximative versions of the SBE, DCE-A, and DCE-B estimator are denoted by $\sim$SBE, $\sim$DCE-A, and $\sim$DCE-B, respectively. Since the independence assumption does not hold in practice, the approximative weights will generate a moderate increase in the mean squared error (MSE) of the CFO estimates.

Please cite this article as: G.J. González, et al., Data-aided CFO estimators based on the averaged cyclic autocorrelation, Signal Processing (2012), http://dx.doi.org/10.1016/j.sigpro.2012.07.032

5.2. Computational complexity

Table 1 summarizes the number of real multiplications in each algorithm and their approximative versions. We assume that a complex multiplication amounts to four real multiplications and a real division is equivalent to one real multiplication.

We note that the approximative algorithms provide a significant reduction in the computational load. Although the algorithm in [12] uses approximately 20% less operations than ∼SBE, it only employs half of the available autocorrelation coefficients ($J/2$). The reduced complexity comes at the price of increased MSE in the CFO estimate, as shown by simulations in the following section. The complexity of DCE-A, DCE-B, SBE, and the estimator in [13] is similar because they all require a matrix inversion.

6. Numerical results

In this section we illustrate the performances of the proposed algorithms and compare them to the Morelli estimator [12] and the Minn estimator [13]. The Minn estimator is implemented using 'Method B' described in [13].

We consider a system with $N=64$ samples and period $M=8$. The length of the cyclic prefix is 16, the channel taps are given by $\{h(l)\}_{l=0}^{L-1}$ with $L=10$, and exponential decay profile $E\{|h(l)|^2\}=Ge^{-l/2}$, where G is chosen such that $\sum_{l=0}^{L-1}E\{|h(l)|^2\}=1$. CFO estimates are averaged over 200 different channel realizations, and each channel realization is averaged over 200 noise realizations for each SNR. The study also considers the CRLB for CFO estimation employing a periodic training sequence over a Rayleigh fading channel derived in [18]. The MLE in (13) is obtained through an exhaustive grid search.

Fig. 1 illustrates the performance of the SBE and the MLE for two different values of the CFO. As expected, the MLE has the best performance. We note that the SBE outperforms the Morelli and the Minn estimators at low SNR. At moderate SNR (higher than 4 dB), the MLE, the SBE and the Minn estimator have similar performance and outperform the Morelli estimator. This is because the Morelli estimator only employs $J/2$ autocorrelation phases. At low SNR (lower than 4 dB), the Minn estimator loses robustness due to the coarse estimation described in Section 2.1 and its performance decreases considerably. All algorithms converge to the CRLB for high SNR (higher than 10 dB). We see that for large CFO (Fig. 1b) the performance of all algorithms decreases except that of the MLE. The reason is that the phase non-linearity (discontinuity in π) causes large errors. The SBE and the estimator in [13] have similar performance at high SNR, outperforming the estimator in [12].

Fig. 2 illustrates the DCE-A estimator for two different values of the CFO. At low SNR (Fig. 2a), we note that the

Table 1
Complexity comparison.

Estimator	No. of real multiplications	$M=J=8$
DCE-A	$(J-1)^3/3+(2M+10)(J-1)^2+(2N+6M+1)(J-1)$	2630
DCE-B	$(J-1)^3/3+10M(J-1)^2+(2N+M)(J-1)+4M$	5020
SBE	$(J-1)^3/3+10(J-1)^2+(J-1)(2N-1)$	1510
Minn [13]	$(J-1)^3/3+2M(J-1)^2+4J(J-1)+4N$	1380
Morelli [12]	$2J(3/4N-M/2)$	700
∼ DCE-A	$(J-1)(2N+6M+2)$	1250
∼ DCE-B	$(J-1)(2N+M+1)+3M$	980
∼ SBE	$2(J-1)(N+1)$	910

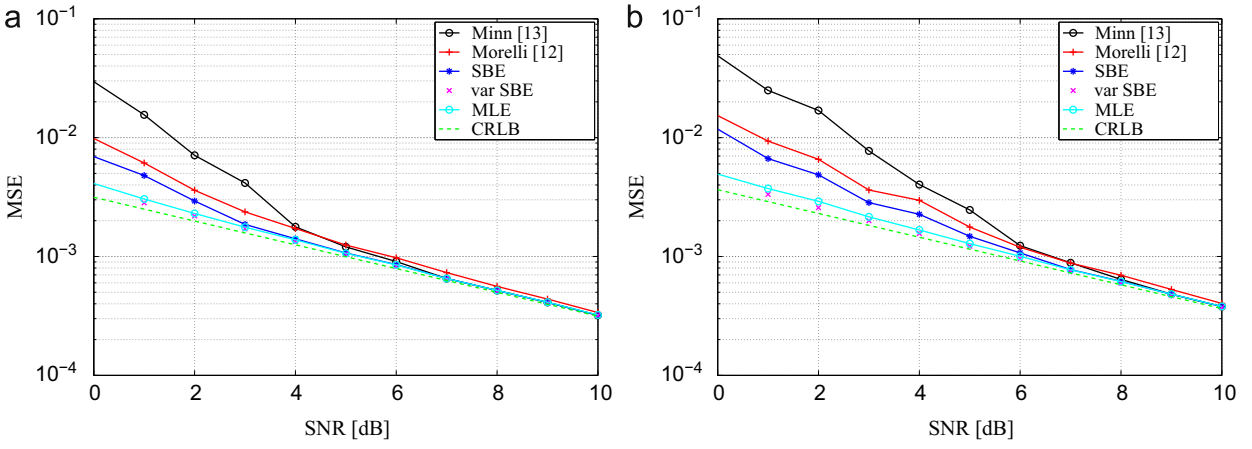

Fig. 1. MSE versus SNR for the SBE for two different CFOs $\epsilon=0.1$ and $\epsilon=0.3$.

Please cite this article as: G.J. González, et al., Data-aided CFO estimators based on the averaged cyclic autocorrelation, Signal Processing (2012), http://dx.doi.org/10.1016/j.sigpro.2012.07.032

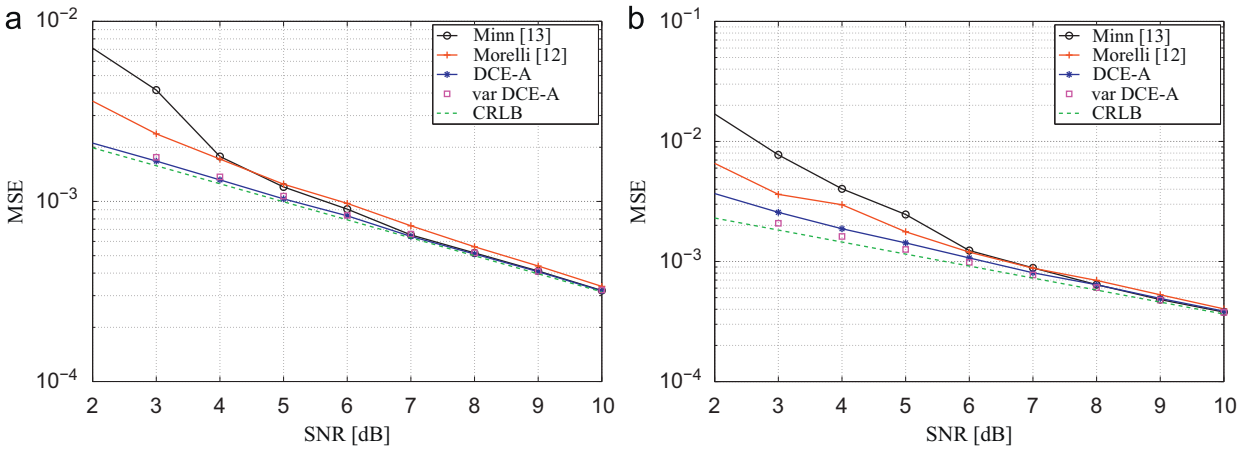

Fig. 2. MSE versus SNR for the DCE-A for two different CFOs $\epsilon = 0.1$ and $\epsilon = 0.3$.

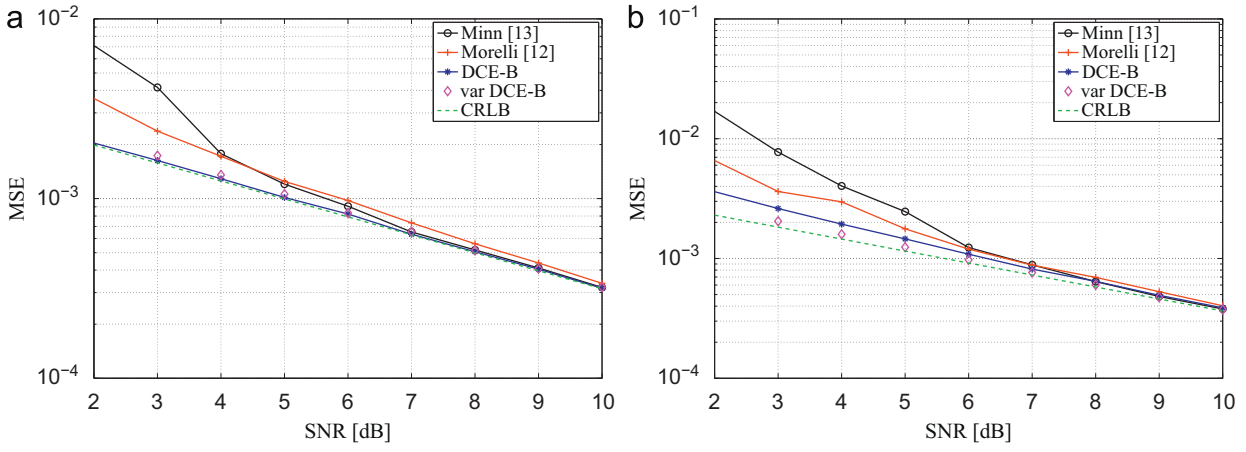

Fig. 3. MSE versus SNR for the DCE-B for two different CFOs $\epsilon = 0.1$ and $\epsilon = 0.3$.

DCE-A algorithm performs better than the Morelli and the Minn estimators. At high SNR, the DCE-A and the Minn estimator have outperform the Morelli estimator. We see that for high CFO (Fig. 2b), the performance of the DCE-A deteriorates because of the large CFO and the phase discontinuity (see DCE-A remarks in Section 4.2). At high SNR values, the DCE-A and the Minn estimators outperform the Morelli estimator.

Fig. 3 illustrates the performance of the DCE-B estimator for two different values of the CFO. The DCE-B works better than the Morelli and the Minn estimators. The comparison with the DCE-B for large CFO is illustrated in Fig. 3b. The large CFO also reduces the performance of the DCE-B estimator, as in the case of the DCE-A. This can be explained by the lack of robustness in the combination over k (see DCE-B remarks in Section 4.2).

Fig. 4 compares the SBE, the DCE-A, and the DCE-B with their approximative versions presented in Section 5.1. We see that the performance of the approximative estimators is slightly degraded, but they still outperform the Morelli and the Minn estimators.

Fig. 5 illustrates the MSE for a longer training sequence with $N=512$ and $M=16$. As can be seen, the difference between algorithms is much smaller, except for the case of the Minn estimator that does not improve the performance at low SNR despite an increasing N.

Covariance matrices $\boldsymbol{\Xi}_{\mathrm{d}}$, $\boldsymbol{\Theta}_p(k)$ and $\boldsymbol{\Theta}_k(p)$ employed in SBE, DCE-A and DCE-B depend on σ^2. The noise variance estimate $\widehat{\sigma^2}$ is typically available in wireless receivers as they have to continuously measure the quality of the communication link. In Fig. 6 we analyze the performance of the algorithms considering an error in the estimation of the noise variance. For the SBE and the Minn estimators, the estimation error does not affect the performance, whereas for the DCE-A and the DCE-B we see a slight degradation.

DCE-based algorithms improve the performance compared with the SBE, but they suffer from a higher bias than SBE at low SNR, because of their poor robustness as discussed in Section 4.2.[1] In summary, the DCE is

[1] In practical wireless communication systems the CFO remains moderate, and the degradation for large CFO of DCE-based estimators should not pose a problem. For example, LTE [19] requires a 1 ppm oscillator accuracy. Together with typical mobile speeds this gives rise to CFO levels in the order of hundreds of Hertz with a subcarrier spacing of 15 kHz.

Please cite this article as: G.J. González, et al., Data-aided CFO estimators based on the averaged cyclic autocorrelation, Signal Processing (2012), http://dx.doi.org/10.1016/j.sigpro.2012.07.032

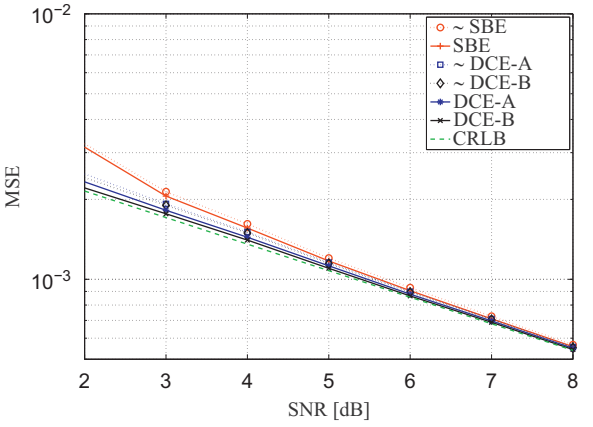

Fig. 4. MSE versus SNR for approximate algorithms for a CFO $\epsilon = 0.1$.

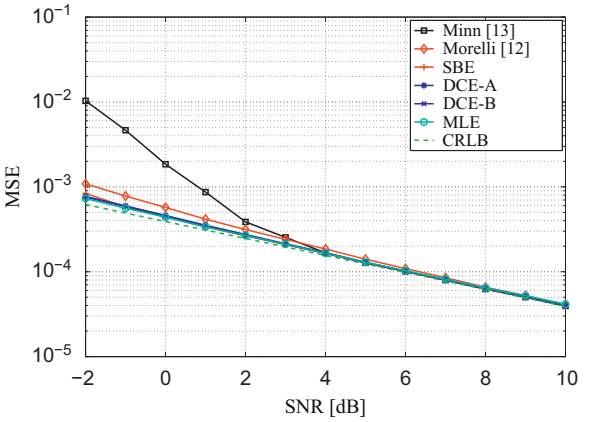

Fig. 5. MSE versus SNR for Morelli [12], Minn [13], the DCE-A, the DCE-B, and the MLE algorithms, for $N=512$ subcarriers, a training sequence of period $M=16$, and a CFO $\epsilon = 0.1$.

preferred for small CFOs and moderate to high SNRs. On the other hand, the SBE performs better than the algorithms in [12,13] for every CFO and SNR.

Table 2 summarizes the gains of the new estimators against the classical ones in terms of the relative MSE. The relative MSE is defined as $MSE_r = Q \times 100/Q_{CR}$, where Q and Q_{CR} are the MSE of the evaluated method and the CRLB at the specified SNR, respectively. The algorithms are evaluated for $\epsilon = 0.1$, $N=64$ and $M=8$.

7. Conclusions

We introduced a family of novel carrier frequency offset (CFO) estimators, a sum-based CFO estimator (SBE) and two direct phase combining CFO estimators (DCE). The algorithms are based on a new interpretation of the cyclic autocorrelation function of the periodic training sequence, referred to as the *averaged cyclic autocorrelation* (ACA). We showed that it is possible to employ all available information in the autocorrelation function when high-order noise terms are included in the formulation of the estimators. The SBE generalizes the classic estimators proposed by Morelli and Mengali and

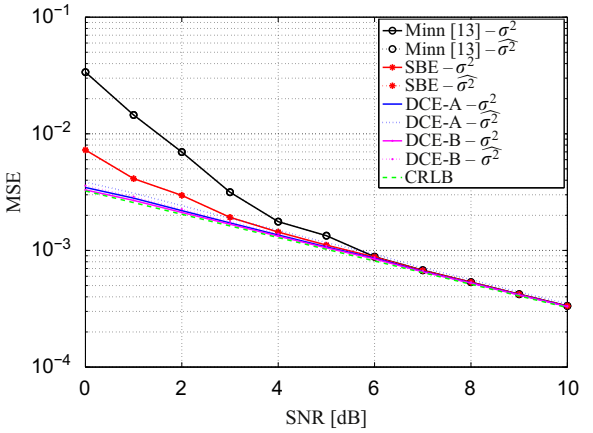

Fig. 6. Performance degradation of the SBE, the DCE-A, the DCE-B, and the algorithm in [13] for an error in the estimation of the noise variance. $\sigma^2/\widehat{\sigma^2} = 10$ where σ^2 is the actual noise variance and $\widehat{\sigma^2}$ the variance estimation.

performs better for low and high SNRs. The DCE algorithms improve the performance as well, but they are sensitive to large CFO. We showed also that the maximum likelihood estimator for CFO can be written in terms of the new autocorrelation function. Moreover, we derived three low-complexity versions of the SBE and DCE algorithms that avoid inversion of the covariance matrix. Although these estimators compromise the performance when compared to SBE and DCE, they still outperform the classical estimators for low CFOs.

Appendix A. Statistics of the cyclic autocorrelation function

Replacing $y(n)$ from (1) into $\hat{r}_c(p,k)$ of (12), we obtain

$$\hat{r}_c(p,k) = \frac{1}{J-k}\sum_{u=0}^{J-k-1}[e^{-j2\pi\epsilon kM/N}|s(p)|^2 + e^{j2\pi\epsilon(p+uM)/N}s(p)w^*(p+(u+k)M) \quad (A.1)$$

$$+e^{-j2\pi\epsilon(p+(u+k)M)/N}s^*(p)w(p+uM) + w(p+uM)w^*(p+(u+k)M)] = e^{-j2\pi\epsilon k/J}|s(p)|^2 + w_c(p,k), \quad (A.2)$$

where $w_c(p,k)$ is a zero-mean correlated noise process for $0 \le p \le M-1$ and $1 \le k \le J-1$. Contrary to the derivation in [3], $w_c(p,k)$ contains the high-order noise term $w(p+uM)w^*(p+(u+k)M)$, which makes (A.1) non-Gaussian.

To obtain CFO estimators based on the ACA, it is necessary to calculate the mean and the covariance matrix of this new function. The mean of (A.1) is given by

$$\mu_c(p,k) = E\{\hat{r}_c(p,k)\} = E\left\{\frac{1}{J-k}\sum_{u=0}^{J-k-1} y(uM+p)y^*((u+k)M+p)\right\} = e^{-j2\pi\epsilon k/jN}|s(p)|^2. \quad (A.3)$$

Since the ACA function is not circularly symmetric Gaussian, it is necessary to consider its real and imaginary components to find the covariance matrix. Starting from the case of fixed p and $1 \le k \le J-1$, we define the

Please cite this article as: G.J. González, et al., Data-aided CFO estimators based on the averaged cyclic autocorrelation, Signal Processing (2012), http://dx.doi.org/10.1016/j.sigpro.2012.07.032

Table 2
Mean squared error relative to the CRLB, defined as $Q \times 100/Q_{CR}$, where Q and Q_{CR} are the MSE of the evaluated method and the CRLB at the specified SNR, respectively.

Estimator	MSE at 0 dB (%)	MSE at 2 dB (%)	MSE at 6 dB (%)	MSE at 10 dB (%)
DCE-A	–	105	104	101
DCE-B	–	104	103	100
SBE	182	122	106	101
Minn [13]	766	227	107	101
Morelli [12]	250	157	120	106

following vectors:

$$\mathbf{r}_k(p) = [\Re\{\tilde{\mathbf{r}}_k^{\mathrm{T}}(p)\}\ \Im\{\tilde{\mathbf{r}}_k^{\mathrm{T}}(p)\}]^{\mathrm{T}}, \tag{A.4}$$

$$\tilde{\mathbf{r}}_k(p) = [\hat{r}_c(p,1)\dots\hat{r}_c(p,J-1)]^{\mathrm{T}},$$

$$\boldsymbol{\mu}_k(p) = [\Re\{\tilde{\boldsymbol{\mu}}_k^{\mathrm{T}}(p)\}\ \Im\{\tilde{\boldsymbol{\mu}}_k^{\mathrm{T}}(p)\}]^{\mathrm{T}}, \tag{A.5}$$

$$\tilde{\boldsymbol{\mu}}_k(p) = [\mu_c(p,1)\dots\mu_c(p,J-1)]^{\mathrm{T}}. \tag{A.6}$$

To obtain an explicit form of the covariance matrix $\mathbf{C}_k(p)$ for $1 \le k \le J-1$ and fixed p, we divide the matrix into four submatrices, as

$$\mathbf{C}_k(p) = E\{(\mathbf{r}_k(p)-\boldsymbol{\mu}_k(p))(\mathbf{r}_k(p)-\boldsymbol{\mu}_k(p))^{\mathrm{T}}\} \tag{A.7}$$

$$= \begin{pmatrix} \boldsymbol{\Sigma}_{11}(p) & \boldsymbol{\Sigma}_{12}(p) \\ \boldsymbol{\Sigma}_{21}(p) & \boldsymbol{\Sigma}_{22}(p) \end{pmatrix}, \tag{A.8}$$

where $\boldsymbol{\Sigma}_{ij}(p)$ are $(J-1)\times(J-1)$ covariance and cross-covariance matrices of the real and imaginary parts of $\tilde{\mathbf{r}}_k(p)$. The components of $[\boldsymbol{\Sigma}_{ij}(p)]_{k_1,k_2}$ are given by

$$[\boldsymbol{\Sigma}_{11}(p)]_{k_1,k_2} = \frac{1}{(J-k_1)(J-k_2)},$$

$$\times \begin{cases} (J-k_1)(\sigma^2|s(p)|^2+\sigma^4/2) & \\ \quad +(J-2k_1)\sigma^2|s(p)|^2 & \\ \quad \cos(4\pi\epsilon k_1/J) & \text{if } k_1=k_2 \text{ and } k_1<J/2, \\ (J-k_1)(\sigma^2|s(p)|^2+\sigma^4/2) & \text{if } k_1=k_2 \text{ and } k_1\ge J/2, \\ \min(J-k_1,J-k_2)(\sigma^2|s(p)|^2 & \\ \quad \cos(2\pi\epsilon(k_2-k_1)/J)+\sigma^4/2) & \\ \quad +(J-k_1-k_2)\sigma^2|s(p)|^2 & \\ \quad \cos(2\pi\epsilon(k_1+k_2)/J) & \text{if } k_1\ne k_2 \text{ and } k_1+k_2<J, \\ \min(J-k_1,J-k_2)(\sigma^2|s(p)|^2 & \\ \quad \cos(2\pi\epsilon(k_2-k_1)/J)+\sigma^4/2) & \text{if } k_1\ne k_2 \text{ and } k_1+k_2\ge J, \end{cases} \tag{A.9}$$

$$[\boldsymbol{\Sigma}_{12}(p)]_{k_1,k_2} = \frac{1}{(J-k_1)(J-k_2)}$$

$$\times \begin{cases} -(J-2k_1)\sigma^2|s(p)|^2\sin(4\pi\epsilon k_1/J) & \text{if } k_1=k_2 \text{ and } k_1<J/2, \\ \min(J-k_1,J-k_2)\sigma^2|s(p)|^2 & \\ \quad \sin(2\pi\epsilon(k_1-k_2)/J) & \\ \quad -(J-k_1-k_2)\sigma^2|s(p)|^2 & \\ \quad \sin(2\pi\epsilon(k_1+k_2)/J) & \text{if } k_1+k_2<J, \\ \min(J-k_1,J-k_2)\sigma^2|s(p)|^2 & \\ \quad \sin(2\pi\epsilon(k_1-k_2)/J) & \text{if } k_1+k_2\ge J, \end{cases} \tag{A.10}$$

$$[\boldsymbol{\Sigma}_{22}(p)]_{k_1,k_2} = \frac{1}{(J-k_1)(J-k_2)}$$

$$\times \begin{cases} (J-k_1)(\sigma^2|s(p)|^2+\sigma^4/2) & \\ \quad -(J-2k_1)\sigma^2|s(p)|^2 & \\ \quad \cos(4\pi\epsilon k_1/J) & \text{if } k_1=k_2 \text{ and } k_1<J/2, \\ (J-k_1)(\sigma^2|s(p)|^2+\sigma^4/2) & \text{if } k_1=k_2 \text{ and } k_1\ge J/2, \\ \min(J-k_1,J-k_2)(\sigma^2|s(p)|^2 & \\ \quad \cos(2\pi\epsilon(k_2-k_1)/J)+\sigma^4/2) & \\ \quad -(J-k_1-k_2)\sigma^2|s(p)|^2 & \\ \quad \cos(2\pi\epsilon(k_1+k_2)/J) & \text{if } k_1\ne k_2 \text{ and } k_1+k_2<J, \\ \min(J-k_1,J-k_2)(\sigma^2|s(p)|^2 & \\ \quad \cos(2\pi\epsilon(k_2-k_1)/J)+\sigma^4/2) & \text{if } k_1\ne k_2 \text{ and } k_1+k_2\ge J, \end{cases} \tag{A.11}$$

and $\boldsymbol{\Sigma}_{21}(p) = \boldsymbol{\Sigma}_{12}^{\mathrm{T}}(p)$.

Consider now the case of k fixed, $k\ne 0$ and $0 \le p \le M-1$. We first define the following vectors:

$$\mathbf{r}_p(k) = [\Re\{\tilde{\mathbf{r}}_p^{\mathrm{T}}(k)\}\ \Im\{\tilde{\mathbf{r}}_p^{\mathrm{T}}(k)\}]^{\mathrm{T}}, \tag{A.12}$$

$$\tilde{\mathbf{r}}_p(k) = [\hat{r}_c(0,k)\dots\hat{r}_c(M-1,k)]^{\mathrm{T}},$$
$$\boldsymbol{\mu}_p(k) = [\Re\{\tilde{\boldsymbol{\mu}}_p^{\mathrm{T}}(k)\}\ \Im\{\tilde{\boldsymbol{\mu}}_p^{\mathrm{T}}(k)\}]^{\mathrm{T}}, \tag{A.13}$$

$$\tilde{\boldsymbol{\mu}}_p(k) = [\mu_c(0,k)\dots\mu_c(M-1,k)]^{\mathrm{T}}.$$

Then, to obtain the covariance matrix $\mathbf{C}_p(k)$ for $0 \le p \le M-1$ and fixed k ($k\ne 0$) the matrix is divided into four submatrices,

$$\mathbf{C}_p(k) = E\{(\mathbf{r}_p(k)-\boldsymbol{\mu}_p(k))(\mathbf{r}_p(k)-\boldsymbol{\mu}_p(k))^{\mathrm{T}}\}, \tag{A.14}$$

$$\mathbf{C}_p(k) = \begin{pmatrix} \boldsymbol{\Delta}_{11}(k) & \boldsymbol{\Delta}_{12}(k) \\ \boldsymbol{\Delta}_{21}(k) & \boldsymbol{\Delta}_{22}(k) \end{pmatrix}, \tag{A.15}$$

where $\boldsymbol{\Delta}_{ij}(k)$ are $(M-1)\times(M-1)$ covariance and cross-covariance matrices of the real and imaginary parts of $\tilde{\mathbf{r}}_p(k)$. Since the ACA terms are uncorrelated for different p, the components of the covariance matrix $\mathbf{C}_p(k)$ can be obtained as

$$[\boldsymbol{\Delta}_{ij}(k)]_{p_1,p_2} = \begin{cases} [\boldsymbol{\Sigma}_{ij}(p_1)]_{k,k} & \text{if } p_1=p_2, \\ 0 & \text{if } p_1\ne p_2. \end{cases} \tag{A.16}$$

For both cases (fixed k or p) the ACA probability density function (pdf) can be approximated by applying the *central limit theorem* to (A.1).[2] As a consequence we

[2] The convergence of the theorem is assured because the terms in (A.1) are independent and identically distributed for a given k.

Please cite this article as: G.J. González, et al., Data-aided CFO estimators based on the averaged cyclic autocorrelation, Signal Processing (2012), http://dx.doi.org/10.1016/j.sigpro.2012.07.032

can establish that $\mathbf{r}_k(p)$ and $\mathbf{r}_p(k)$ are joint Gaussian vectors of mean $\boldsymbol{\mu}_k(p)$ and $\boldsymbol{\mu}_p(k)$; and covariance matrix $\mathbf{C}_k(p)$ or $\mathbf{C}_p(k)$, respectively. The accuracy of the approximation depends on the values J and k.

A discussion on the optimal value of J considering the channel condition is presented in [18], where it is shown that for modern multicarrier communication systems (LTE, WiMAX [9,14,10]) J is close to 20 and never less than 10. For small values of k the statistical approximation is good, but as k increases the approximation becomes coarse. Nevertheless, we stress that the derivation of the estimators based on the ACA only require to know the covariance matrix, which is obtained in the closed form.

Appendix B. Covariance matrix of SBE weighting factors

The procedure to obtain the covariance matrix $\boldsymbol{\Xi}_{\mathrm{d}} = E\{(\boldsymbol{\xi}_{\mathrm{d}} - E\{\boldsymbol{\xi}_{\mathrm{d}}\})(\boldsymbol{\xi}_{\mathrm{d}} - E\{\boldsymbol{\xi}_{\mathrm{d}}\})^{\mathrm{T}}\}$ is the following:

1. We derive the covariance matrix of the ACA terms summed over p, defined as
$$r_{\mathrm{s}}(k) = \sum_{p=0}^{M-1} \hat{r}_{\mathrm{c}}(p,k) = e^{-j2\pi\epsilon k/J} S + w_{\mathrm{s}}(k) \tag{B.1}$$
where $S = \sum_{p=0}^{M-1} |s(p)|^2$ and $w_{\mathrm{s}}(k) = \sum_{p=0}^{M-1} w(p,k)$.
2. Assuming high SNR, the covariance matrix of the phase $\xi(k) = \arg\{r_{\mathrm{s}}(k)\}$ is approximated by a matrix transformation, from the covariance matrix derived in 1.
3. Finally, a second transformation that define the difference of consecutive $\xi(k)$ is used to derive $\boldsymbol{\Xi}_{\mathrm{d}}$.

Since the ACA terms are uncorrelated for different p, the covariance matrix of $r_{\mathrm{s}}(k)$, $\mathbf{S}_k$, can be obtained from $[\boldsymbol{\Sigma}_{ij}(p)]_{k_1,k_2}$ as

$$\mathbf{S}_k = \begin{pmatrix} \boldsymbol{\Omega}_{11} & \boldsymbol{\Omega}_{12} \\ \boldsymbol{\Omega}_{21} & \boldsymbol{\Omega}_{22} \end{pmatrix} \tag{B.2}$$

with components

$$[\boldsymbol{\Omega}_{ij}]_{k_1,k_2} = \sum_{p=0}^{M-1} [\boldsymbol{\Sigma}_{ij}(p)]_{k_1,k_2} \tag{B.3}$$

where $[\boldsymbol{\Sigma}_{ij}(p)]_{k_1,k_2}$ follows from (A.9) to (A.11) of Appendix A. As $w_{\mathrm{s}}(k)$ is not conjugate symmetric, the covariance matrix $\mathbf{S}_k$ is also $2(J-1) \times 2(J-1)$.

In the high SNR region, the phase of $r_{\mathrm{s}}(k)$ can be obtained as

$$\begin{aligned}
\xi(k) &= \arg\{e^{-j2\pi\epsilon k/J} S + w_{\mathrm{s}}(k)\} \\
&= \frac{-2\pi\epsilon k}{J} + \arctan\left\{\frac{\Im\{w_{\mathrm{s}}(k) e^{j2\pi\epsilon k/J}\}}{S + \Re\{w_{\mathrm{s}}(k) e^{j2\pi\epsilon k/J}\}}\right\} \\
&\approx \frac{-2\pi\epsilon k}{J} + \frac{\Im\{w_{\mathrm{s}}(k) e^{j2\pi\epsilon k/J}\}}{S + \Re\{w_{\mathrm{s}}(k) e^{j2\pi\epsilon k/J}\}} \\
&\approx \frac{-2\pi\epsilon k}{J} + \frac{\Im\{w_{\mathrm{s}}(k) e^{j2\pi\epsilon k/J}\}}{S} = \frac{-2\pi\epsilon k}{J} + w_{\xi}(k),
\end{aligned} \tag{B.4}$$

where $w_{\xi}(k) = \Im\{w_{\mathrm{s}}(k) e^{j2\pi\epsilon k/J}\}/S$. From the relation between $w_{\xi}(k)$ and $w_{\mathrm{s}}(k)$, we can define a linear transform that allow us to obtain the covariance matrix of $\xi(k)$ from $\mathbf{S}_k$ in (B.2). The transformation can be represented by the following matrix:

$$\begin{aligned}
\mathbf{U}_k = \frac{1}{S}\Bigg[&\mathrm{diag}\left[\sin\left(\frac{2\pi\epsilon}{J}\right), \ldots, \sin\left(\frac{2\pi\epsilon(J-1)}{J}\right)\right] \\
&\mathrm{diag}\left[\cos\left(\frac{2\pi\epsilon}{J}\right), \ldots, \cos\left(\frac{2\pi\epsilon(J-1)}{J}\right)\right]\Bigg].
\end{aligned} \tag{B.5}$$

Then, using the affine property, the covariance matrix for $\xi(k)$, $\boldsymbol{\Xi}_k$, has the following form:

$$\boldsymbol{\Xi}_k = \mathbf{U}_k \mathbf{S}_k \mathbf{U}_k^{\mathrm{T}}, \tag{B.6}$$

with the components given by

$$\begin{aligned}
[\boldsymbol{\Xi}]_{k_1,k_2} = &\frac{\sigma^2}{S(J-k_1)(J-k_2)} \\
&\times \begin{cases}
(J-k_1)(1+M\sigma^2/(2S)) - (J-2k_1) & \text{if } k_1 = k_2 \text{ and } k_1 < J/2, \\
(J-k_1)(1+M\sigma^2/(2S)) & \text{if } k_1 = k_2 \text{ and } k_1 \geq J/2, \\
\min(J-k_1, J-k_2) - (J-k_1-k_2) & \text{if } k_1 \neq k_2 \text{ and } k_1 + k_2 < J, \\
\min(J-k_1, J-k_2) & \text{if } k_1 = k_2 \text{ and } k_1 + k_2 \geq J.
\end{cases}
\end{aligned} \tag{B.7}$$

Note that $\boldsymbol{\Xi}_k$ no longer depends on ϵ. The matrix transformation that defines the differences of $\xi(k)$ can be described as

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 1 \end{bmatrix}. \tag{B.8}$$

Now it is possible to obtain the covariance matrix of $\boldsymbol{\xi}_{\mathrm{d}}$, $\boldsymbol{\Xi}_{\mathrm{d}}(k)$, from $\boldsymbol{\Xi}_k$ as follows:

$$\boldsymbol{\Xi}_{\mathrm{d}} = \mathbf{T} \boldsymbol{\Xi}_k \mathbf{T}^{\mathrm{T}}. \tag{B.9}$$

Appendix C. Covariance of $\alpha(p,k)$

Since $\alpha(p,k)$ is the argument of the ACA function, it is a function of two-variables. In the following we derive first the covariance matrix of $\alpha(p,k)$ for $0 \leq p \leq M-1$ considering k fixed ($\boldsymbol{\Theta}_p(k)$), and then, the covariance matrix of $\alpha(p,k)$ for $1 \leq k \leq J-1$ with p fixed ($\boldsymbol{\Theta}_k(p)$). The procedure followed to obtain the covariance matrices is similar to the derivation in Appendix B.

From (15), we have at high SNR

$$\begin{aligned}
\alpha(p,k) &= \arg\{e^{-j2\pi\epsilon k/J} |s(p)|^2 + w_{\mathrm{c}}(p,k)\} \\
&= \frac{-2\pi\epsilon k}{J} + \arctan\left\{\frac{\Im\{w_{\mathrm{c}}(p,k) e^{j2\pi\epsilon k/J}\}}{|s(p)|^2 + \Re\{w_{\mathrm{c}}(p,k) e^{j2\pi\epsilon k/J}\}}\right\} \\
&\approx \frac{-2\pi\epsilon k}{J} + \frac{\Im\{w_{\mathrm{c}}(p,k) e^{j2\pi\epsilon k/J}\}}{|s(p)|^2 + \Re\{w_{\mathrm{c}}(p,k) e^{j2\pi\epsilon k/J}\}} \\
&\approx \frac{-2\pi\epsilon k}{J} + \frac{\Im\{w_{\mathrm{c}}(p,k) e^{j2\pi\epsilon k/J}\}}{|s(p)|^2} \\
&\approx \frac{-2\pi\epsilon k}{J} + w_{\alpha}(p,k),
\end{aligned} \tag{C.1}$$

where $w_{\alpha}(p,k) = \Im\{w_{\mathrm{c}}(p,k) e^{j2\pi\epsilon k/J}\}/|s(p)|^2$ is a zero-mean correlated noise process for $0 \leq p \leq M-1$ and $1 \leq k \leq J-1$ (as in (21)). From the relation between $w_{\alpha}(p,k)$ and $w_{\mathrm{c}}(p,k)$, it is possible to define a linear transform that allow us to obtain the covariance matrix $\boldsymbol{\Theta}_p(k)$ for fixed k.

Please cite this article as: G.J. González, et al., Data-aided CFO estimators based on the averaged cyclic autocorrelation, Signal Processing (2012), http://dx.doi.org/10.1016/j.sigpro.2012.07.032

The transformation can be represented in matrix form as

$$\mathbf{V}_p(k)=\left[\operatorname{diag}\left[\frac{\sin\left(\frac{2\pi\epsilon k}{J}\right)}{|s(1)|^2},\ldots,\frac{\sin\left(\frac{2\pi\epsilon k}{J}\right)}{|s(M-1)|^2}\right]\right.$$
$$\left.\operatorname{diag}\left[\frac{\cos\left(\frac{2\pi\epsilon k}{J}\right)}{|s(1)|^2},\ldots,\frac{\cos\left(\frac{2\pi\epsilon k}{J}\right)}{|s(M-1)|^2}\right]\right]. \tag{C.2}$$

Finally, the expression of the covariance matrix is

$$\mathbf{\Theta}_p(k)=\mathbf{V}_p(k)\mathbf{C}_p(k)\mathbf{V}_p(k)^{\mathrm{T}}, \tag{C.3}$$

whose components, taking into account (A.8) of Appendix A, are given by

$$[\mathbf{\Theta}_p(k)]_{p_1,p_2}=\frac{\sigma^2}{(J-k)^2}$$
$$\times\begin{cases}(J-k)(1+\sigma^2/(2|s(p)|^2))-(J-2k))/|s(p)|^2 & \text{if } p_1=p_2 \text{ and } k<J/2,\\ (J-k)(1+\sigma^2/(2|s(p)|^2))/|s(p)|^2 & \text{if } p_1=p_2 \text{ and } k\geq J/2,\\ 0 & \text{c.c.}\end{cases} \tag{C.4}$$

Similarly, considering now p fixed in (C.1), the matrix transformation to obtain $\mathbf{\Theta}_k(p)$ is defined as

$$\mathbf{V}_k(p)=\frac{1}{|s(p)|^2}\left[\operatorname{diag}\left[\sin\left(\frac{2\pi\epsilon}{J}\right),\ldots,\sin\left(\frac{2\pi\epsilon(J-1)}{J}\right)\right]\right.$$
$$\left.\operatorname{diag}\left[\cos\left(\frac{2\pi\epsilon}{J}\right),\ldots,\cos\left(\frac{2\pi\epsilon(J-1)}{J}\right)\right]\right]. \tag{C.5}$$

Then, the resulting covariance matrix is

$$\mathbf{\Theta}_k(p)=\mathbf{V}_k(p)\mathbf{C}_k(p)\mathbf{V}_k(p)^{\mathrm{T}}, \tag{C.6}$$

where, taking into account (A.15) of Appendix A, the terms result

$$[\mathbf{\Theta}_k(p)]_{k_1,k_2}=\frac{\sigma^2}{|s(p)|^2(J-k_1)(J-k_2)}$$
$$\times\begin{cases}(J-k_1)(1+\sigma^2/(2|s(p)|^2))-(J-2k_1) & \text{if } k_1=k_2 \text{ and } k_1<J/2,\\ (J-k_1)(1+\sigma^2/(2|s(p)|^2)) & \text{if } k_1=k_2 \text{ and } k_1\geq J/2,\\ \min(J-k_1,J-k_2)-(J-k_1-k_2) & \text{if } k_1\neq k_2 \text{ and } k_1+k_2<J,\\ \min(J-k_1,J-k_2) & \text{if } k_1\neq k_2 \text{ and } k_1+k_2\geq J.\end{cases} \tag{C.7}$$

Here we stress that derivation of DCE only requires the covariance matrices (C.3) and (C.6), that are obtained in closed form. If more statistical information of $\alpha(p,k)$ is needed, then we can approximate its pdf from (A.2) using the matrix transformation defined in (C.1), after applying the central limit theorem. Then, it is possible to establish that $w_\alpha(p,k)$ is a real joint Gaussian process, for $0\leq p\leq M-1$ and $1\leq k\leq J-1$, and zero mean. The covariance matrices of $w_\alpha(p,k)$ for fixed k or p are defined respectively in (C.3) and (C.6).

Appendix D. Bias of autocorrelation-based estimators for the low SNR region

We show here that CFO estimators based on the phase differences are biased at low SNR regime. The bias comes from the phase discontinuity at π. CFO estimation algorithms based on the phase of the autocorrelation function employ the BLUE to linearly combine available data and obtain the estimation. A general expression for the CFO estimate $\hat{\epsilon}$ is

$$\hat{\epsilon}=\sum_{k=1}^{J-1}a_g(k)\beta_L^{(d)}(k), \tag{D.1}$$

where $a_g(k)$ are the BLUE weighting factors and $\beta_L^{(d)}(k)$ are the phase differences of the autocorrelation function (see Sections 2.1 and 4). The linear model assumes that $\beta_L^{(d)}(k)$ does not have discontinuities in π, i.e., it is non-ambiguous. If the autocorrelation phase circumscribed to the interval $(-\pi,\pi]$ is denoted as $\beta_C^{(d)}(k)$; the relation between $\beta_L^{(d)}(k)$ and $\beta_C^{(d)}(k)$ can be expressed as

$$\beta_L^{(d)}(k)=\beta_C^{(d)}(k)+2\pi G(k), \tag{D.2}$$

where $G(k)\in\{-1,0,1\}$ as it is necessary to suppress the ambiguity in $\beta_C^{(d)}(k)$.[3] As $\beta_L^{(d)}(k)$ is not available, in (D.1) it is employed $\beta_C^{(d)}(k)$, then

$$\hat{\epsilon}=\sum_{k=1}^{J-1}a_g(k)\{\beta_L^{(d)}(k)-2\pi G(k)\}$$
$$=\sum_{k=1}^{J-1}a_g(k)\beta_L^{(d)}(k)-2\pi\sum_{k=1}^{J-1}a_g(k)G(k). \tag{D.3}$$

If the CFO is low and the SNR is high for some realization, $G(k)=0$ with high probability. Alternatively, if the CFO is large (close to $J/2$) and/or the SNR is low, it is highly probable that $G(k)\in\{-1,0\}$ or $G(k)\in\{0,1\}$.[4] Taking into account (D.3) the bias in the CFO estimation is given by

$$E\{\hat{\epsilon}\}-\epsilon=E\left\{\sum_{k=1}^{J-1}a_g(k)\beta_L^{(d)}(k)\right\}-\epsilon-E\left\{2\pi\sum_{k=1}^{J-1}a_g(k)G(k)\right\}$$
$$=\left(\frac{-J}{2\pi}\right)\left(\frac{-2\pi\epsilon}{J}\right)-\epsilon-2\pi\sum_{k=1}^{J-1}E\{a_g(k)G(k)\},$$
$$=-2\pi\sum_{k=1}^{J-1}E\{a_g(k)G(k)\}, \tag{D.4}$$

where it is assumed that $E\{\beta_L^{(d)}(k)\}=-2\pi\epsilon/J$ as in (B.4) and $\sum_{k=1}^{J-1}a_g(k)=-J/2\pi$. From (D.4) we can infer that the bias is not zero if $G(k)=1$ or $G(k)=-1$ for some k. Then, the CFO estimate is biased with high probability for high CFO or low SNR.

In previous derivation it is assumed that $E\{\beta_L^{(d)}(k)\}=-2\pi\epsilon/J$, what it is not evident for low SNR. In the following it is included a verification of this assumption for the phase corresponding to the SBE, employing for simplicity the Gaussian approximation introduced in Appendix A. It must be noted that the phases used in SBE and [12] are equivalent, as discussed in Section 4.1.

The general expression of the phase differences is $\beta_L^{(d)}(k)=\beta(k)-\beta(k-1)$ with $\beta(0)=0$. Then if $E\{\beta(k)\}=0$, $E\{\beta_L^{(d)}(k)\}=0$. A general expression of $\beta(k)$ is

$$\beta(k)=\arg\{e^{-j2\pi\epsilon k/J}D(k)+\mathcal{V}(k)\}$$

[3] E.g. if $\beta_C^{(d)}(k)=-2,-3,3$, respectively for $k=1,2,3$, then $G(k)=0,0,-1$; resulting in $\beta_L^{(d)}(k)=-2,-3,-3,28$.

[4] It is impossible that $G(k)$ takes values 1 and -1 in the same realization, for different k.

Please cite this article as: G.J. González, et al., Data-aided CFO estimators based on the averaged cyclic autocorrelation, Signal Processing (2012), http://dx.doi.org/10.1016/j.sigpro.2012.07.032

$$=\frac{-2\pi\epsilon k}{J}+\arctan\left\{\frac{\Im\{\mathcal{V}(k)e^{j2\pi\epsilon k/J}\}}{D(k)+\Re\{\mathcal{V}(k)e^{j2\pi\epsilon k/J}\}}\right\}, \quad (D.5)$$

where $D(k)$ refers to the signal energy term and $\mathcal{V}(k)$ is the noise term (including the signal–noise cross products), as can be seen in (B.4), (C.1) and [12,13]. The usual high SNR assumption disregards the noise term in the denominator. Alternatively, in the following we study the phase (D.5) considering that: (1) SNR $\to 0$, i.e. the noise term $\mathcal{V}(k)$ is dominant over the signal term $D(k)$ and (2) J and k are large enough to assume that $\mathcal{V}(k)$ is approximately Gaussian, as it is specified in Appendix A.

The outline of the proof is the following: (1) we show that the argument of the arctangent in (D.5) has Cauchy distribution, (2) it is derived that the probability density function of the arctangent of the Cauchy distribution, denoted as $\hat{w}_\beta(k)$, and (3) it is shown that the mean of $\hat{w}_\beta(k)$ is zero.

Taking into account previous considerations, (D.5) can be written as

$$\beta(k)\approx\frac{-2\pi\epsilon k}{J}+\arctan\left\{\frac{w_{1(k)}}{w_{2(k)}}\right\}. \quad (D.6)$$

In case of SBE (and the algorithm proposed in [12]), the expressions of $w_1(k)$ and $w_2(k)$ are

$$w_1(k)=\Re\{w_s(k)\}\sin\left(\frac{2\pi\epsilon k}{J}\right)+\Im\{w_s(k)\}\cos\left(\frac{2\pi\epsilon k}{J}\right), \quad (D.7)$$

$$w_{2(k)}=\Re\{w_s(k)\}\cos\left(\frac{2\pi\epsilon k}{J}\right)-\Im\{w_s(k)\}\sin\left(\frac{2\pi\epsilon k}{J}\right), \quad (D.8)$$

where $w_s(k)$ is defined in Appendix B. When $\sigma_1(k)$ and $\sigma_2(k)$ denote the standard deviation of $w_1(k)$ and $w_2(k)$, respectively, and $\rho(k)=E\{w_1(k)w_2(k)\}/(\sigma_1(k)\sigma_2(k))$ is the normalized correlation, the random variable $w_1(k)/w_2(k)\to\mathbf{C}(\mu_c(k),\gamma_c(k))$, where $\mathbf{C}$ is the Cauchy distribution of median $\mu_c(k)=\rho(k)\sigma_1(k)/\sigma_2(k)$ and scale parameter $\gamma_c(k)=\sigma_1(k)\sqrt{1-\rho(k)^2}/\sigma_2(k)$.

The probability density function of $\hat{w}_\beta(k)=\arctan\{w_1(k)/w_2(k)\}$ is

$$f_w(x)=\frac{\sec(x)^2}{\pi\gamma_c((\tan(x)-\mu_c)^2/\gamma_c^2+1)}. \quad (D.9)$$

It is easy to show that the mean of $\hat{w}_\beta(k)$ is non-zero if $\mu_c(k)$ is non-zero, i.e., if $w_1(k)$ and $w_2(k)$ are correlated. The expression of $\rho(k)$ in terms of the covariance matrices is

$$\rho(k)=\frac{1}{\sigma_1(k)\sigma_2(k)}\left[\frac{1}{2}\sin\left(\frac{4\pi\epsilon k}{J}\right)([\mathbf{\Omega}_{11}]_{k,k}-[\mathbf{\Omega}_{22}]_{k,k})+\cos\left(\frac{4\pi\epsilon k}{J}\right)[\mathbf{\Omega}_{12}]_{k,k}\right]=0, \quad (D.10)$$

where $[\mathbf{\Omega}_{ij}]_{k_1,k_2}$ is defined in (B.2). Then $\hat{w}_\beta(k)$ has zero mean and, as a consequence, $\beta_L^{(d)}(k)$ has zero mean.

References

[1] T. Pollet, M. Van Bladel, M. Moeneclaey, BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise, IEEE Transactions on Communications 43 (1995) 191–193.
[2] T. Schmidl, D. Cox, Robust frequency and timing synchronization for OFDM, IEEE Transactions on Communications 45 (1997) 1613–1621.
[3] M. Morelli, U. Mengali, Carrier-frequency estimation for transmissions over selective channels, IEEE Transactions on Communications 48 (2000) 1580–1589.
[4] M. Ghogho, A. Swami, Unified framework for a class of frequency-offset estimation techniques for OFDM, in: IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), vol. 4, pp. iv-361–iv–364.
[5] W. Gardner, Cyclostationarity in Communications and Signal Processing, IEEE Press, 1994.
[6] F. Gini, G. Giannakis, Frequency offset and symbol timing recovery in flat-fading channels: a cyclostationary approach, IEEE Transactions on Communications 46 (1998) 400–411.
[7] M. Ghogho, A. Swami, T. Durrani,On blind carrier recovery in time-selective fading channels, in: Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers, 1999, vol. 1, pp. 243–247.
[8] H. Bölcskei, Blind estimation of symbol timing and carrier frequency offset in wireless OFDM systems, IEEE Transactions on Communications 49 (2001) 988–999.
[9] E. Dahlman, S. Parkvall, J. Sköld, P. Beming, 3G Evolution HSPA and LTE for Movile Broadband, 2nd ed. Elsevier, 2008.
[10] IEEE. 802.16, Part 16: Air interface for fixed and mobile broadband wireless access systems, amendment 2: Physical and medium access control layers for combined fixed and mobile operation in licensed bands and corrigendum 1, December 2005.
[11] M.-o. Pun, M. Morelli, C.C.J. Kuo, Multi-Carrier Techniques For Broadband wireless Communications: A Signal Processing Perspectives, Imperial College Press, London, UK, 2007.
[12] M. Morelli, U. Mengali, An improved frequency offset estimator for OFDM applications, IEEE Communications Letters 3 (1999) 75–77.
[13] H. Minn, P. Tarasak, V. Bhargava, OFDM frequency offset estimation based on BLUE principle, in: IEEE Vehicular Technology Conference (VTC) (2002), vol. 2, pp. 1230–1234.
[14] H. Holma, A. Toskala, LTE for UMTS-OFDMA and SC-FDMA Based Radio Access, 1st ed. Wiley, 2009.
[15] S. Barbarossa, M. Pompili, G. Giannakis, Channel-independent synchronization of orthogonal frequency division multiple access systems, IEEE Journal on Selected Areas in Communications 20 (2002) 474–486.
[16] M. Ghogho, A. Swami,G. Giannakis, Optimized null-subcarrier selection for CFO estimation in OFDM over frequency-selective fading channels, in: Proceedings of the IEEE GLOBECOM, pp. 202–206.
[17] S. Kay, Fundamentals of statistical Signal Processing, Volume I: Estimation Theory, 1st ed., Prentice Hall, PTR, 1993.
[18] M. Ghogho, P. Ciblat, A. Swami, P. Bianchi, Training design for repetitive-slot-based CFO estimation in OFDM, IEEE Transactions on Signal Processing 57 (2009) 4958–4964.
[19] G.T. 36.101, 3rd generation partnership project; technical specification group radio access network; evolved universal terrestrial radio access (E-UTRA); user equipment (UE) radio transmission and reception.

Please cite this article as: G.J. González, et al., Data-aided CFO estimators based on the averaged cyclic autocorrelation, Signal Processing (2012), http://dx.doi.org/10.1016/j.sigpro.2012.07.032