

E-reference

SYNCHRONIZATION

Mounir Ghogho^{1,2}, *Philippe Ciblat*³ and *Ananthram Swami*⁴

¹School of Electronic and Electrical Engineering, University of Leeds, UK

²International University of Rabat, Morocco

³Telecom Paris-Tech, Paris, France

⁴US Army Research Lab, Maryland, USA

Key Words: synchronization, OFDM(A), timing offset, phase offset, carrier frequency offset, data-aided, blind, training design.

Abstract:

1. Introduction

Synchronization is fundamental to the proper operation of wireless sensor networks (WSNs) and mobile ad hoc networks (MANETs). Synchronization may be performed at various levels: from very coarse to very fine-grained; the required accuracy depends upon the application and the network environment, and hence can change dynamically. For example, coarse time synchronization suffices for many detection and tracking applications, but finer sync is required for distributed array processing, analog to digital conversion (ADC), slot sync, and cooperative communications.

Network synchronization is a well-studied topic with an extensive history, especially for wired networks, e.g., see Lindsey et al. [1], Bregni [2], and references therein. Typically, these works assumed high quality devices, availability of fine control of the network, extensive connectivity with little or no mutual interference, as well as often assuming known (or repeatable and measurable) propagation and processing delays [3]. Surveys of WSN sync protocols may be found in the papers by Sivrikaya and Yener [4], Johannessen [5], Sundararaman et al. [6], [3], and [7].

According to Wiki, synchronization is timekeeping which requires the coordination of events to operate a system in unison. The familiar conductor of an orchestra serves to keep the orchestra in time. Systems operating with all their parts in synchrony are said to be synchronous or in sync. Some systems may be only approximately synchronized, or pleiochronous. For some applications relative offsets between events need to be determined, for others only the order of the event is important.

Wireless broadcast and sync are much older problems, of course. A fascinating book by Peter Galison [8] describes the pioneering work of Einstein and Poincare on finding common time references (circa 1909). Albert Einstein, then a young, obscure German physicist was experimenting with measuring time using telegraph networks and with the coordination of clocks at train stations; the renowned mathematician Henri Poincaré, president of the French Bureau of Longitude, was mapping time coordinates across continents. Sundials and watches have been in use for centuries; they were relatively inaccurate; but travel and communications were slow (until at least the end of the nineteenth century), and those local time differences were of little importance. Story of Huygens, 1665 and his observation of syncing pendulums while on the sickbed

Even older of course are time-keeping mechanisms in nature; Stogatz [9] provides an elegant description of synchronization among fireflies in Malaysia, and circadian rhythms. Studies of the robustness of circadian clocks (e.g., via clock neurons in the superchiasmatic nucleus) indicate that individual neurons are sloppy timekeepers but synchronized neurons are precise clocks. could lead to development of algorithms for synchronizing communication networks (e.g., transmit beamforming, GPS) that exploit combinations of local and global signaling [10].

Synchronization is crucial in diverse applications:

- In energy constrained WSNs and MANETs, accurate clock synchronization facilitates energy-efficient Medium Access Control (MAC).
- Reliable coherent communication systems require accurate synchronization (timing and frequency) and channel estimation, particularly when the data rate or bandwidth

is large, and in multiple antenna systems. Blind estimation techniques have been proposed in order to save bandwidth; however, the tradeoff with power and complexity is ambiguous. Practical systems typically use some form of training which may be prepended, appended, or embedded in the data packet. Training typically leads to low complexity receivers and good performance with moderate sacrifice in rate.

- Synchronization is crucial to enabling distributed communications such as cooperative communications and relaying, and in network management (probing and monitoring).
- Many networked signal processing tasks depend critically upon a common time reference; examples include event detection and target tracking, localization, (multi-modal) sensor fusion, distributed array processing, and synchronization of distributed information caches.
- In robotics, synchronization and delay management are crucial components of networked control and actuation; poor synchronization can lead to control instabilities. Other applications in collaborative robotics such as mapping or geolocation, also require a common time reference.
- WSN setup and maintenance also require synchronization. Node geolocation is typically required after deployment, and this, in turn, requires time synchronization.

We have alluded only to the time synchronization problem in the preceding paragraphs. But frequency synchronization is equally crucial. The ensuing sections in this chapter will study the synchronization problem in detail. Our focus is on systems based on Orthogonal Frequency Domain Access (OFDM). We will consider both the single and multi-user problem, and the impact of flat-fading as well as frequency-selective channels. In these sections, the focus is on time and frequency synchronization with a base station. In the last chapter we will consider the network time synchronization problem, without the constraint (or freedom) of OFDM.

The chapter is written at the beginner graduate level. The reader is assumed to have a working knowledge of digital communications (e.g., [11], [12]), estimation theory (e.g., [13]).

2. Synchronization for flat fading channel

In this section, we assume that the multipath delay spread in the propagation channel does not lead to Inter-Symbol Interference (ISI). In practice, this holds if the symbol period is sufficiently larger than the delay spread of the channel. The frequency response of the channel is thus flat. As the channel does not induce ISI, the OFDM modulation scheme does not provide any advantages and a (old-fashioned) single-carrier approach suffices.

From a practical point of view, such an assumption on the flat fading channel is satisfied in satellite communications (e.g., DVB-S), anisotropic transmissions (e.g., between two DVB-T transmitters), optical fiber communications (e.g., with pre-compensation of the static chromatic dispersion) [14, 15].

The transmitted signal (in baseband) [11, 12] is given by

$$x_a(t) = \sum_{n=0}^{N-1} s_n h_a(t - kT_s)$$

where

- s_n are the transmitted symbols belonging to ASK, PSK or QAM constellations or any linearly precoded scheme. N is the sequence length used for estimating the synchronization (sync) parameters;
- T_s is the symbol period which is assumed to be known throughout this chapter. For more information about the blind estimation of T_s , the reader is referred to [16];
- $h_a(t)$ is the shaping filter that classically is a square-root Nyquist filter. For example, in 3G, it is the square-root raised cosine filter with roll-off 0.35. The main property that we need further is that the filter is band-limited with bandwidth between $1/T_s$ and $2/T_s$.

For the flat fading channel, the received signal (in baseband) is

$$y_a(t) = x_a(t - \tau_0) e^{2\pi i(f_0 t + \phi_0)} + b_a(t)$$

where

- the sync parameters are the symbol timing τ_0 , the (constant) phase ϕ_0 , and the carrier frequency offset (CFO) f_0 [17, 18, 19].
- $b_a(t)$ is the complex-valued circularly-symmetric white zero-mean Gaussian process with variance N_0 per real dimension (for more details about the assumption on the noise, see [12]). Note that the concept of circularity is of great importance and will be defined in detail later.

The received signal can be re-written as follows

$$y_a(t) = \left(\sum_{n=0}^{N-1} s_n h_a(t - kT_s - \tau_0) \right) e^{2i\pi(f_0 t + \phi_0)} + b_a(t)$$

We would like to point out that we do not assume an *a priori* distribution for the timing and phase parameters. Indeed, the timing can belong uniformly into the interval $[0, T_s)$ since the receiver and the transmitter are not synchronized in time yet. Due to the channel propagation, the phase can take any value over $[0, 2\pi)$. In contrast, the CFO which is due either to local oscillator mismatch or Doppler effects can only lie in a pre-defined interval. To illustrate this let us consider the worst case when only a cheap local oscillator is available. Its precision is about 40ppm which leads to a CFO of 40 kHz at carrier frequency 1GHz. Assuming a rather high vehicle speed of 360km/h, the Doppler induced offset is upper-bounded by 333Hz. Consequently, the CFO is much smaller than the typical signal bandwidth which is of order MHz. Note that the main source of the CFO is the local oscillator mismatch and not the Doppler effect. Indeed, the Doppler effect has a greater

influence on the coherence time of the channel and thus on the statistical model of the channel (is it a fast flat fading channel or a slow flat fading channel?). The nature of the flat fading channel (fast or slow) is crucial for designing properly the communication scheme (feedback link, diversity management, etc.) but not the synchronization step since, except in very infrequent cases, the synchronization step duration is always much smaller than the channel coherence time.

Before going further, we recap the optimal symbol detector when synchronization is perfect ($\tau_0 = 0$, $\phi_0 = 0$, and $f_0 = 0$) [11, 12, 18, 17]. If the information symbols are equally likely, the maximum likelihood detector is the optimal one in the sense of error probability minimization. Therefore, we have

$$\{\hat{s}_n\}_{n=0}^{N-1} = \arg \max_{\{s_n\}_{n=0, \dots, N-1}} p(y(t) | \{s_n\}_{n=0, \dots, N-1}).$$

Due to the Gaussianness of the noise process, we have [20]

$$p(y(t) | \{s_n\}_{n=0, \dots, N-1}) \propto e^{-\int_{\mathbb{R}} |y(t) - \sum_{n=0}^{N-1} s_n h(t - nT_s)|^2 dt / 2N_0}.$$

Thus

$$\hat{s}_n = \arg \min_{\{s_n\}_{n=0, \dots, N-1}} \int_{\mathbb{R}} \left| y(t) - \sum_{n=0}^{N-1} s_n h(t - nT_s) \right|^2 dt$$

We finally obtain that

$$\hat{s}_n = \arg \min_{\{s_n\}_{n=0, \dots, N-1}} J_N(s)$$

where

$$J_N(s) = \int_{\mathbb{R}} |y(t)|^2 dt + \sum_{n, n'=0}^{N-1} s_n s_{n'}^* \tilde{h}(n - n') - 2 \sum_{n=0}^{N-1} \Re\{s_n^* z(n)\}$$

where

- $z_a(t) = h_a(-t) \star y_a(t)$ the continuous-time output of the so-called matched filter,
- $z(n) = z_a(nT_s)$ the sampled version (at symbol rate) of the matched filter output,
- $\tilde{h}_a(t) = h_a(-t) \star h_a(t)$ the continuous-time equivalent filter,
- and $\tilde{h}(n) = \tilde{h}_a(nT_s)$ the discrete equivalent filter.

We remark that the optimality criterion depends on the received signal only through the discrete-time matched filter output. The introduction of the signal $z(n)$ can be also justified by following an alternative way: the received signal (under the perfect synchronization assumption), can be re-written as follows:

$$y_a(t) = \sum_n s_n \Phi_n(t) + b_a(t)$$

where $\Phi_n(t) = h_a(t - nT_s)$. As the useful information in $y_a(t)$ is the symbols s_n , the received signal can be split into two parts: the useful one associated with the signal subspace spanned by the functions $\Phi_n(t)$, and the other one generated by the space orthogonal to the

signal subspace, the so-called "noise subspace". Let $\tilde{z}(n) = \langle y_a(t) | \Phi_n(t) \rangle$ be the projected signal onto the signal subspace where $\langle \bullet | \bullet \rangle$ denotes the canonical inner product. One can easily check that $\tilde{z}(n) = z(n)$. As the noise is Gaussian, its contribution to the signal subspace is independent of those to the orthogonal noise subspace. Therefore, without loss of optimality, one needs to only work with $\tilde{z}(n) = z(n)$ and can drop the projection onto the noise subspace.

Moreover, one can easily see that

$$z(n) = \sum_k \tilde{h}(k) s_{n-k} + \tilde{b}(n) \quad (1)$$

where $\tilde{b}(n) = h_a(-t) \star b_a(t)|_{t=nT_s}$ is circularly-symmetric Gaussian noise with zero mean and power spectral density $S_{\tilde{b}}(e^{2i\pi f}) = 2N_0 \tilde{h}(e^{2i\pi f}) = 2N_0 \sum_k \tilde{h}(k) e^{-2i\pi k f}$. The maximum likelihood (ML) criterion depends on the shaping and propagation filters only through the so-called discrete-time equivalent filter \tilde{h} . Therefore the system performance will only be driven by the filter \tilde{h} and the SNR.

Now the second step of the optimal detector is to find the minimum of $J_N(s)$. When the shaping filter $h_a(t)$ is a square-root Nyquist filter, it is well known that $\tilde{h}_a(t)$ is the Nyquist filter and $\tilde{h}(n) = \delta_{0,n}$ where $\delta_{0,n}$ is the Kronecker index. Thus the function $J_N(s)$ can be significantly simplified to

$$J_N(s) = \sum_{n=0}^{N-1} |z(n) - s_n|^2.$$

Consequently, the optimal detector is a symbol-by-symbol detector

$$\hat{s}_n = \arg \min_{s_n} |z(n) - s_n|^2$$

which is the so-called threshold detector. When the Nyquist condition is not satisfied (especially when the channel is non flat fading), the minimization of $J_N(s)$ is much harder and can be done via the famous Viterbi algorithm [21, 22]. When the Viterbi algorithm is too complex (channel too long and/or high constellation size) suboptimal detectors, such as the zero-forcing (ZF), minimum mean-square error (MMSE), decision-feedback equalizer (DFE), can be used. For details, mathematical explanations and derivations, we refer the reader to [11, 12]. The optimal receiver is summarized in Fig.1.

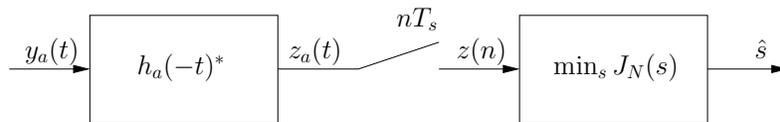


Figure 1: Optimal receiver structure (with perfect synchronization)

When the sync parameters are non-zero but known, the continuous-time received signal is given by

$$y_a(t) = \sum_n s_n \Psi_n(t - nT_s) + b_a(t)$$

with $\Psi_n(t) = h_a(t - nT_s - \tau_0)e^{2i\pi(f_0t + \phi_0)}$. Once again, the optimal operation at the receiver side is to project $y_a(t)$ onto the function $\Psi_n(t)$. Therefore the optimal receiver is now given in Fig.2.

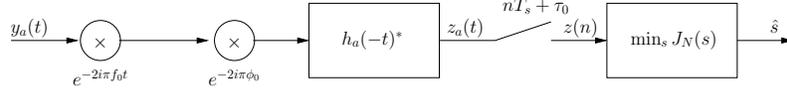


Figure 2: Optimal receiver structure (with known synchronization parameters)

Notice that the phase compensator can be located anywhere in Fig.2 since it commutes with the other operators. The CFO compensator can be located anywhere iff the CFO is small enough compared to the filter bandwidth¹. This last assumption is usually satisfied as mentioned earlier.

Let us now move on to some performance evaluation when the synchronization step is not carried out. In order to understand the influence of each parameter, we will consider parameters one by one, assuming all the others are known. Let us first focus on the timing. In the flat fading context, the shaping filter is usually square-root Nyquist and more precisely a square-root raised cosine filter with roll-off ρ . Under the perfect synchronization assumption, $z(n)$ is not distorted by Inter-Symbol Interference (ISI). But if the timing is not perfectly known, $z(n)$ will be affected by ISI. In Fig. 3 (left), we display the eye diagram of $z_a(t)$ when $\tau_0 = 0$ and $\rho = 0.5$ with BPSK modulation. We remark that if the sampling operation is not done at a multiple of T_s , the eye will be less open and performance will be degraded. In Fig. 3 (right), bit-error rate (BER) versus SNR E_b/N_0 curves are shown for different values of τ_0 . Notice that performance degrades significantly when the timing error exceeds 10% of the symbol period.

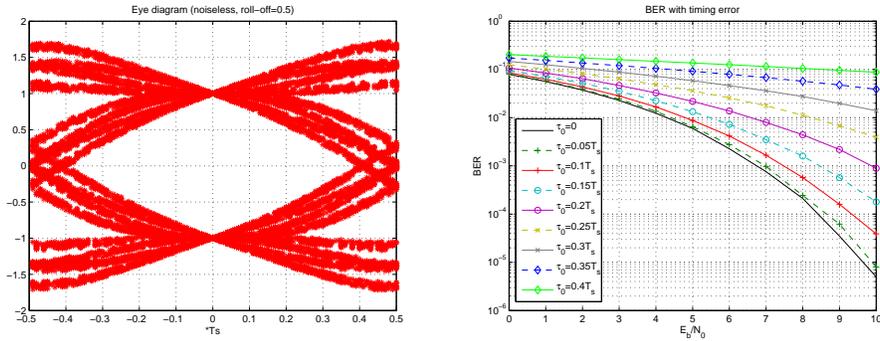


Figure 3: Eye diagram (left) in noiseless case and BER (right) versus E_b/N_0 in presence of timing mis-synchronization.

¹Let $h_1(t)$ and $h_2(t)$ be two filters of bandwidth $1/T_s$. Let us assume also that $f_0T_s \ll 1$. We have $h_1(t) \star (h_2(t)e^{2i\pi f_0 t}) = e^{2i\pi f_0 t} \int H_2(v)H_1(v + f_0)e^{2i\pi vt} dv = e^{2i\pi f_0 t} (h_1(t) \star h_2(t)) + o(f_0)$ where H_k is the Fourier Transform of h_k . The second equality holds since f_0 is small compared to $1/T_s$. We thus conclude that the CFO operation can be permuted with the filtering operator.

For Fig. 4,

We next consider perfect timing synchronization, but without any phase synchronization. The shaping filter is once again a square-root raised cosine filter with roll-off $\rho = 0.5$. In Fig. 4 (left), we plot the samples $z(n)$ (so before decision) when BPSK is employed and the phase shift is equal to 0.1 at $E_b/N_0 = 10\text{dB}$. The constellation is thus rotated and the disk is now closer to the decision threshold which induces an increase of the BER. In Fig. 4 (right), we display the BER versus E_b/N_0 for different values of phase shift. Performance degradation is significant.

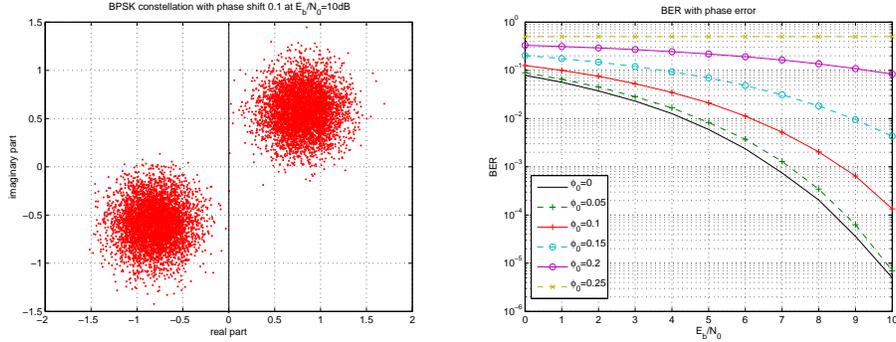


Figure 4: Samples $z(n)$ before decision (left) when $\phi_0 = 0.1$ and $E_b/N_0 = 10\text{dB}$ and BER (right) versus E_b/N_0 in the presence of phase mis-synchronization.

We next examine the influence of the CFO on performance. The same simulation set-up is used as previously. In Fig. 5 (left), we plot the samples $z(n)$ when the CFO is equal to 0.01 at $E_b/N_0 = 30\text{dB}$. We remark that the BPSK constellation is rotated with different rotation angles $f_0 n$ at each time index n which leads to a circle if the frame is long enough. Due to the noise, we observe a ring. In Fig. 5 (right), we compute the BER versus E_b/N_0 for different values of CFO. The frame length in this example is 1000 data symbols.

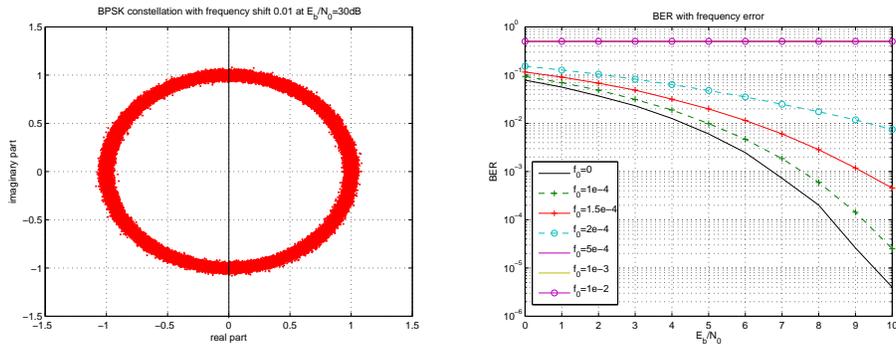


Figure 5: Samples $z(n)$ before decision (left) when $f_0 = 0.01$ and $E_b/N_0 = 30\text{dB}$ and BER (right) versus E_b/N_0 in the presence of frequency mis-synchronization.

For each synchronization parameter, we saw a loss in performance that leads to com-

munication failure. Therefore, we clearly need to add a synchronization step at the receiver side in order to keep the performance as high as possible.

As we just saw, synchronization parameters have a great influence on system performance, so that estimating these parameters is crucial:

- Why should we estimate the timing parameter τ_0 : Knowledge of τ_0 is crucial for choosing the correct sampling time when the continuous-time signal is converted to the (symbol rate) discrete-time signal (see Fig. 3). The need for sampling at the symbol rate at the true timing offset has its origin on the non-satisfaction of the Nyquist-Shannon sampling theorem. Therefore, by oversampling at least at the baud rate (but in practice twice the symbol rate), the discrete-time signal $\check{y}(n) = y_a(nT_s/2)$ contains all the information in $y_a(t)$. Then the “symbol timing” matching can be done via digital processing, namely, interpolation. More interestingly, by considering the bivariate process $\mathbf{y}(n) = [y_a(nT_s), y_a(nT_s + T_s/2)]^T$ which corresponds exactly to $\check{y}(n)$, a fractionally-spaced (FS) equalizer can be used to retrieve the information symbol. To implement this, one needs to know the filters $h_1(n) = h_a(nT_s)$ and $h_2(n) = h_a(nT_s + T_s/2)$. Thus the symbol timing estimation issue has been replaced by a pair of filter estimation issues. Notice that at least one filter (usually $h_2(n)$ does not satisfy the Nyquist criterion) is non-flat fading in its discrete-time version. Another way to cope with timing offset is to incorporate the timing into the channel by rewriting the received signal as:

$$y_a(t) = \left(\sum_{n=0}^{N-1} s_n \check{h}_a(t - nT_s) \right) e^{2i\pi(f_0 t + \phi_0)} + b_a(t)$$

where

$$\check{h}_a(t) = h_a(t - \tau_0)$$

is the new equivalent channel. Now the channel is unknown since τ_0 is unknown. Therefore once again the timing estimation issue boils down to a channel estimation issue.

- As the phase rotation is a linear time-invariant operation, it can be viewed as a filtering operator and thus incorporated into the filter. Therefore we have

$$y_a(t) = \left(\sum_{n=0}^{N-1} s_n \mathring{h}_a(t - nT_s) \right) e^{2i\pi f_0 t} + b_a(t)$$

where

$$\mathring{h}_a(t) = h_a(t - \tau_0) e^{2i\pi \phi_0}$$

and the phase estimation issue can be avoided.

- The CFO cannot be viewed as the modification of a linear filter since the CFO is not a time-invariant transformation. Therefore the CFO estimation issue can not be avoided.

In the case of a flat fading channel, if channel estimation is used to synchronize in time and phase (in order to avoid the “real” synchronization steps), we do not use all of our knowledge about the signal, and in particular, its Nyquist property. Nevertheless, if the synchronization step has to be avoided, the estimation issue deals with the joint frequency and (non-flat) channel estimation issue. In the Data-aided mode, such a problem will be studied in Section ???. In the Non-data-aided mode, the CFO can be estimated regardless of the channel, as shown in Section 2.2., and the channel can also be estimated by following the approaches given in “Channel Estimation” E-reference.

The communications protocol may permit the transmitter to explicitly adjust its signalling so as to facilitate receiver side synchronization. Therefore, there are two main classes of estimation problems:

- DA (Data-Aided)²: the receiver knows a sequence of transmitted symbols s_n . These symbols belong to the so-called “training sequence”. Obviously, this leads to a loss in spectral efficiency since during the transmission of the training symbols, no information symbols are sent. Nevertheless this approach (which is very popular in civilian applications such as GSM, 3G, DVB-S) has several advantages since it enables us to dramatically simplify the design and the implementation of the synchronization parameter estimators as will be seen later. Moreover, good performance can usually be attained with only a few (training) symbols, and thus the cost in spectral efficiency remains quite low and acceptable.
- NDA (Non-Data-aided)³: the transmitter does not send any training sequences per se; thus the receiver does not have deterministic knowledge about a symbol subsequence. However, the receiver will have some structural or statistical information about the symbols such as the nature of the constellation, the correlation between the symbols, etc. Such a scenario obviously occurs in passive listening in security applications. In civilian applications, the NDA approach is sometimes useful for tracking parameter fluctuations or in broadcasting applications (e.g., TV). Indeed, the TV application differs completely from a peer-to-peer application (such as GSM between the mobile and the Base Station) since the TV receiver can be switched on at any moment and will not warn the broadcast transmitter that it is switched on (If it does, the broadcast transmitter will always be “interrupted”, since there is no dedicated channel for learnign the channel and synchronization parameters). So the broadcast transmitter will not transmit a training sequence as soon as a TV receiver goes live in the network. To speed up the process, the TV receiver has to synchronize itself in a blindly manner.

Some remarks before going further:

- i) the main property here compared to the non-flat fading channel is that the channel is assumed to be known. The fact that the filter is also a squared-root Nyquist is absolutely not required but will greatly simplified some derivations and thus the algorithms.

²also called, training approach, supervised mode, ...

³also called, blind, unsupervised, ...

- ii) the estimate that we will develop can be used and adapted (sometimes) in the non-flat fading channel case.
- iii) the trade-off between the estimation quality (provided by training) and the remaining time for transmitting data can be studied by the framework of the information theory. A Shannon capacity can be expressed by taking into account the channel uncertainty. The whole system capacity is out of the scope of this tutorial but has been only a few analyzed in the literature for channel estimation in [23].

2.1. DA case

Let $\hat{\theta}_N$ be an unbiased estimator of the sought parameter θ relying on N observation samples. Based on [24,20], we know that the Mean Square Error (MSE), defined as $\mathbb{E}[\|\hat{\theta} - \theta\|^2]$ of any unbiased estimator of θ is lower-bounded by the so-called Cramer-Rao bound (CRB) which will be described mathematically later in this chapter. An estimator whose MSE is equal to the CRB is called *efficient*. Efficient estimators do not exist for many estimation problems. Therefore the notion of asymptotic efficiency has been introduced. This means that the ratio between the MSE and the CRB tends to 1 as the number of samples N goes to infinity. Under mild conditions (given in [24, 20]), the Maximum-Likelihood estimator is asymptotically efficient and it is asymptotically normal.

In view of the asymptotic efficiency and normality, it is natural to consider the maximum-likelihood estimator first. If the ML estimator can be implemented, one often considers the problem to be closed. In contrast, if the ML estimator cannot be implemented (because of its complexity or sometimes even intractability), the estimation issue is open and other estimators have to be found. We hasten to add that if an efficient estimator does not exist, then it may be possible to find better estimators than the ML. Second, the CRB may not be tight when the number of samples (or SNR) is low, and tighter bounds may (such as the Bhattacharya and Ziv-Zakai bounds) may need to be considered. We refer the reader to [24] for details.

So let us start with the introduction of the ML estimator for the joint synchronization parameters. The Likelihood can be written as follows

$$p(y(t)|\tau, \phi, f) \propto e^{-\int_{\mathbb{R}} |y(t) - \sum_{n=0}^{N-1} s_n h_a(t - nT_s - \tau) e^{2i\pi(f t + \phi)}|^2 dt / 2N_0}.$$

where the training sequence, $\{s_n\}$, is known.

The ML estimator is obtained as follows

$$[\hat{\tau}_N, \hat{\phi}_N, \hat{f}_N] = \arg \min_{\tau, \phi, f} \int_{\mathbb{R}} \left| y(t) - \sum_{n=0}^{N-1} s_n h_a(t - nT_s - \tau) e^{2i\pi(f t + \phi)} \right|^2 dt$$

Setting the derivative to zero, and assuming that the CFO is small compared to the bandwidth, we obtain the following set of equations:

$$\begin{cases} \Re\{\sum_{n=0}^{N-1} s_n^* e^{-2i\pi\phi} e^{-2i\pi f T_s n} z_{\tau}^*(n)\} & = 0 \\ \Im\{\sum_{n=0}^{N-1} s_n^* e^{-2i\pi\phi} e^{-2i\pi f T_s n} z_{\tau}(n)\} & = 0 \\ \Im\{\sum_{n=0}^{N-1} s_n^* e^{-2i\pi\phi} n e^{-2i\pi f T_s n} z_{\tau}(n)\} & = 0 \end{cases} \quad (2)$$

where

$$z_\tau(n) = \int_{\mathbb{R}} y_a(t) h_a(t - nT_s - \tau)^* dt = h_a(-t)^* \star y_a(t)|_{t=nT_s+\tau}$$

$$z'_\tau(n) = \int_{\mathbb{R}} y_a(t) h'_a(t - nT_s - \tau)^* dt = h'_a(-t)^* \star y_a(t)|_{t=nT_s+\tau}$$

with $h'_a(t)$ is the derivative of $h_a(t)$.

A typical approach is variable elimination: we try to express a parameter in terms of the others, and eliminate it by using the found expression. Here this algebraic manipulation can be applied to the phase (as a function of the timing and the CFO), thanks to the second line in Eq. (2). But the timing and the CFO cannot then be written as explicit functions due to the non-linearity of these equations. Therefore the joint (timing, phase, CFO) problem is intractable.

We remark that joint frequency and phase estimators can be developed when the timing is known. Indeed, as we will see later, if the timing is known (and thus assumed to be zero without loss of generality), then the second and third equations can lead to practical estimators with reasonable computational complexity. Therefore, we split our problem into two different problems that will be treated separately : i) the timing issue and ii) the phase and CFO estimation issues.

The ordering of the two problems has a great influence on the nature of estimators that will be used, as we see next.

- **First scheme:** Timing is estimated first and then the phase and CFO are estimated, assuming that the timing estimate is perfect. As the DA timing cannot be derived in closed-form when the phase and CFO are unknown, we need to develop a NDA timing estimator that is insensitive to the actual phase and CFO values. The second part of this scheme deals with phase and CFO estimations issues. As the timing is now known, DA estimators can be developed. If training is not available, NDA estimators can also be considered.
- **Second scheme:** the phase and the CFO are estimated first; timing is then estimated, assuming that the phase and CFO estimates are perfect, and thus perfectly corrected. Once again, the phase and CFO estimators here have to be NDA and insensitive to timing error. As the timing can be wrong, the sampled filter can generate Inter-Symbol Interference. Therefore, we need to design joint phase and CFO NDA estimators that can work even when the (non-flat) fading channel is unknown. Assuming that phase and CFO are perfectly compensated, the timing estimator can be blind or aided by a training sequence.

The two schemes are summarized in Fig. 6.

Consequently, we need to solve the following issues (even if training sequences are assumed available!):

- **Problem 1:** DA Phase and CFO estimation (when timing is known and so can be considered to be zero, without loss of generality (wlog)).
- **Problem 1':** DA Timing estimation (when Phase and CFO are known and so can be assumed to be zero wlog).

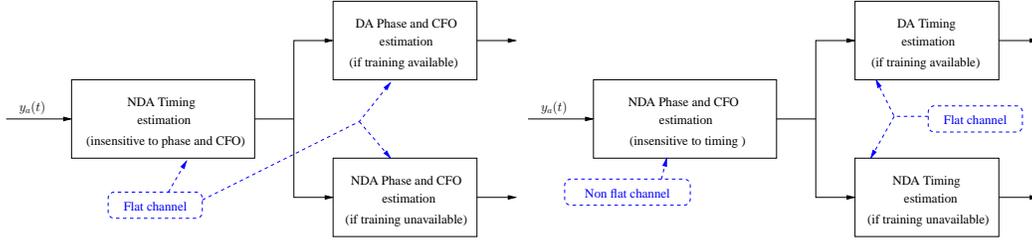


Figure 6: Summary of the first scheme (left) and the second scheme (right).

- **Problem 2:** NDA Phase and CFO estimation (insensitive to unknown timing and hence a non-flat channel),
- **Problem 2':** NDA Timing estimation (insensitive to unknown phase and CFO).

Notice that the first scheme (associated with Problems 1 and 2') is used frequently and advocated in a lot of practical systems. Hereafter, we first focus on Problem 1 and Problem 1' relying on training sequences.

DA timing estimation issue

In this paragraph, we focus on the following estimation problem

$$y_a(t) = \sum_{n=0}^{N-1} s_n h_a(t - nT_s - \tau_0) + b_a(t).$$

After simple algebraic manipulations (similar to those associated with the maximum-likelihood detector), the ML timing estimator can be obtained as follows

$$\hat{\tau}_N = \arg \max_{\tau} \underbrace{\sum_{n=0}^{N-1} \Re\{s_n^* z_{\tau}(n)\}}_{J_N(\tau)} \quad (3)$$

where we recall that $z_{\tau}(n)$ is the output of the matched filter sampled at $nT_s + \tau$.

The cost function $J_N(\cdot)$ of (3) is plotted in Fig. 7 for $E_b/N_0 = 5\text{dB}$ (left) and $E_b/N_0 = 10\text{dB}$ (right) with $N = 100$. The shaping filter is a square-root raised cosine with roll-off $\rho = 0.5$. The value of the sought τ_0 is $0.1T_s$ and the x-axis has been normalized by T_s .

Observe that the cost function is concave around the true point. Therefore one can proceed to find the maximum of $J_N(\cdot)$ in two steps.

- a *coarse* search through a 1-D grid which provides a first estimate $\hat{\tau}_N^{(0)}$ of τ_0 .
- once the coarse search has roughly localized the maximum, one can use a gradient-descent algorithm on the function $J_N(\cdot)$ initialized by $\hat{\tau}_N^{(0)}$, as follows

$$\begin{aligned} \hat{\tau}_N^{(m)} &= \hat{\tau}_N^{(m-1)} + \mu J'_N(\hat{\tau}_N^{(m-1)}) \\ &= \hat{\tau}_N^{(m-1)} + \mu \sum_{n=0}^{N-1} \Re \left\{ s_n^* \frac{\partial z_{\tau}(n)}{\partial \tau} \Big|_{\hat{\tau}_N^{(m-1)}} \right\} \end{aligned} \quad (4)$$

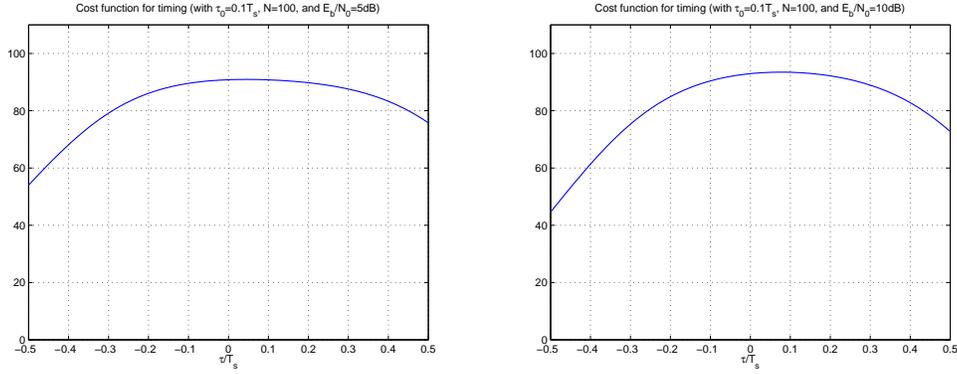


Figure 7: Cost function $J_N(\tau)$ for $E_b/N_0 = 5\text{dB}$ (left) and $E_b/N_0 = 10\text{dB}$ (right) with $N = 100$ and $\tau_0 = 0.1T_s$.

where m is the iteration index of the gradient algorithm. The implementation of the derivative term is actually quite easy since it corresponds to the sampled version (at $\hat{\tau}_N^{(m-1)}$) of the output of the received signal passed through the matched filter associated with the derivative filter $h'_a(t)$. This can be done digitally (via interpolation) if the received signal $y_a(t)$ has been sampled at twice the symbol rate or faster.

In order to benchmark different estimators associated with the same estimation issue, it is useful to have a performance lower-bound on the MSE. The CRB is the most popular one because it is quite simple to derive (especially in DA mode) and tight (we remind that under mild conditions, the ML is asymptotically efficient). Recall that the CRB definition in the context of an unbiased estimator of the timing.

$$\text{CRB}(\tau) = \frac{1}{\mathbb{E} \left[\left(\frac{\partial \log p(y|\tau)}{\partial \tau} \Big|_{\tau_0} \right)^2 \right]} = - \frac{1}{\mathbb{E} \left[\frac{\partial^2 \log p(y|\tau)}{(\partial \tau)^2} \Big|_{\tau_0} \right]}$$

After standard algebraic manipulations, we have

$$\log p(y|\tau) = \frac{1}{N_0} \Re \left\{ \sum_{n=0}^{N-1} s_n^* \int_{\mathbb{R}} y_a(t) h_a(t - nT_s - \tau)^* dt \right\} + \text{cste}$$

Thus we obtain

$$\frac{\partial^2 \log p(y|\tau)}{(\partial \tau)^2} = \frac{1}{N_0} \Re \left\{ \sum_{n=0}^{N-1} s_n^* \int_{\mathbb{R}} y_a(t) h_a''(t - nT_s - \tau)^* dt \right\}$$

with $h_a''(t)$ the second derivative function of $h_a(t)$. By taking the mathematical expectation (and using the zero-mean property of the noise), we obtain

$$\mathbb{E} \left[\frac{\partial^2 \log p(y|\tau)}{(\partial \tau)^2} \right] = \frac{1}{N_0} \Re \left\{ \sum_{n,n'=0}^{N-1} s_n^* s_{n'} \int_{\mathbb{R}} h_a(t - n'T_s) h_a''(t - nT_s - \tau)^* dt \right\}$$

Using Parseval's identity, we finally obtain

$$\mathbb{E} \left[\frac{\partial^2 \log p(y|\tau)}{(\partial \tau)^2} \right] = -\frac{4\pi^2}{N_0} \Re \left\{ \sum_{n,n'=0}^{N-1} s_n^* s_{n'} \int_{\mathbb{R}} f^2 |H(f)|^2 e^{2i\pi(n'-n)fT_s} df \right\}$$

where $H(f)$ is the Fourier transform of $h_a(t)$.

Consequently, the CRB for the timing estimation issue is given by

$$\text{CRB}(\tau) = \frac{N_0}{4\pi^2 \Re \left\{ \sum_{n,n'=0}^{N-1} s_n^* s_{n'} \int_{\mathbb{R}} f^2 |H(f)|^2 e^{2i\pi(n'-n)fT_s} df \right\}}$$

This expression shows the influence of the shaping filter through the integral of the square of $fH(f)$. But the influence of the training sequence (especially its size) is still quite unclear. To address this issue, we will model our training sequence as a realization of a random process. More precisely, s_n is a realization of pseudo-noise stationary process. In practice, any training sequence is generated through a shift register. Let us consider $r_s(m) = \mathbb{E}[s_{n+m}s_n^*]$ the autocorrelation function. and the associated spectrum $S_s(e^{2i\pi f}) = \sum_m r_s(m)e^{-2i\pi m f}$. By using some results on Cesaro sums, one can prove that

$$\frac{1}{N} \sum_{n,n'=0}^{N-1} \mathbb{E}[s_{n'}s_n^*] e^{2i\pi(n'-n)fT_s} \xrightarrow{a.s.} S_s(e^{-2i\pi fT_s}), \text{ when } N \rightarrow \infty.$$

As a consequence, we have

$$\text{CRB}(\tau) = \frac{N_0}{4\pi^2 N \int_{\mathbb{R}} f^2 |H(f)|^2 S_s(e^{-2i\pi fT_s}) df}$$

For channel estimation (when synchronization is perfectly done), a white training sequence (i.e., with flat spectrum) is optimal [25, 26]. But a white sequence is not necessarily the best choice for the synchronization parameters⁴. Further, the best training sequence correlation property may be different for different synchronization parameters. Since a white sequence always leads to reasonable performance, it provides a good trade-off. Therefore, a white training sequence is chosen in current real-life systems.

When the training sequence is white, the CRB simplifies to

$$\text{CRB}(\tau) = \frac{N_0}{4\pi^2 E_s N \int_{\mathbb{R}} f^2 |H(f)|^2 df}$$

where E_s is the variance of s_n .

The CRB provides us some insights about the behavior of the estimates.

- The CRB is $\mathcal{O}(1/N)$. This is logical since the CRB associated with channel estimation (cf. "Channel estimation" E-reference) is also $\mathcal{O}(1/N)$. We recall that the timing could be incorporated into the channel estimation box without loss in performance. Thus fortunately the CRB offers the same behavior.

⁴for instance, for the timing, the best sequence is one whose spectrum is $S_s(e^{2i\pi fT_s}) = \sigma_s^2 \delta((f - f_{\max})T_s)$ where $f_{\max} = \arg \max_f f^2 |H(f)|^2$.

- The CRB is $\mathcal{O}(1/\text{SNR})$. Once again, this result is consistent with that one associated with channel estimation.
- The influence of the shaping filter can be analyzed, especially the influence of the roll-off. Recall that the system is most sensitive to timing error if the roll-off is small and thus if the occupied extra bandwidth is small. Unfortunately, the quality of estimation also decreases when the roll-off becomes small (see [18, 17] for more details).

In Fig. 8, we plot MSE and CRB versus E_b/N_0 and N with BPSK modulation. Unless otherwise stated, $E_b/N_0 = 10\text{dB}$, $N = 100$, and $\rho = 0.5$.

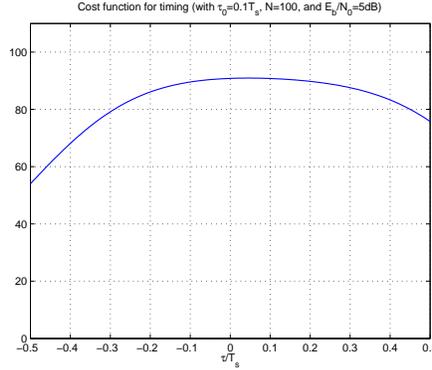


Figure 8: **MSE and CRB versus E_b/N_0 (left) and N (right).**

So far we have considered a blockwise approach. In the past, adaptive approaches (*i.e.*, sample-by-sample) were of great interest due to buffer size limitation and computational complexity. Today, the adaptive approach is mainly useful for tracking parameter variations during data (and not training) transmission. The parameter update can be implemented under two different modes: i) the NDA mode since it is carried out during data transmission, and ii) the Decision-Directed (DD) mode. In the NDA mode, different updates can be developed and usually are obtained through the adaptive version of the blockwise approaches described in Section 2.2.. In the DD mode, different updates are obtained by calculating the adaptive version of the DA estimators and then replacing the training s_n with a decision on data symbol \hat{s}_n . Therefore it is useful to develop an adaptive version of the previously-described DA ML timing estimator. Thus instead of working block-by-block, we work sample-by-sample. The (stochastic) gradient-descent algorithm at time n can be derived from the blockwise version of the gradient descent algorithm (cf. Eq. (4)) by keeping only the derivative term associated with time n . We thus have

$$\hat{\tau}_n = \hat{\tau}_{n-1} + \mu \underbrace{\Re \left\{ s_n^* \frac{\partial z_\tau(n)}{\partial \tau} \right\}}_{e(n)} \Big|_{\hat{\tau}_N^{(m-1)}} \quad (5)$$

where $\hat{\tau}_n$ is the estimated value of τ_0 at time n . Now the time index and the iteration

number coincide. Such an adaptive algorithm can be fully analyzed by using the stochastic approximation tool [27], but this analysis is out of the scope of this chapter.

In DD mode, the symbol s_n which is unknown during data transmission has to be replaced with the (hard) decision of the symbol denoted by \hat{s}_n . DD algorithm can be also applied with soft decision on the symbol as was done in [28]. More details about the DD algorithm with soft decision will be given later, in the context of phase estimation. To perform well, the error probability should be small enough to limit the impact of error propagation. Therefore DD mode is always implemented after an initialization step (feasible thanks to the training sequence) in order to ensure low data detection error. In order to avoid the calculation of the derivative at each symbol period (notice that in the blockwise approach the number of iteration and thus of derivative calculation is much smaller than the number of samples), we replace it with

$$\left. \frac{\partial z_\tau(n)}{\partial \tau} \right|_{\hat{\tau}_N^{(m-1)}} = \frac{z_{\hat{\tau}_N^{(m-1)} + \Delta}(n) - z_{\hat{\tau}_N^{(m-1)} - \Delta}(n)}{2\Delta}$$

where Δ is a design parameter that must be carefully adjusted. We have just described the very popular early-late (adaptive) estimator [29]. Notice that other update equations with *ad hoc* $e(n)$ have been proposed in the literature [30, 31] and often perform much better.

DA phase and CFO estimation issue

In this paragraph, we assume that the timing is known and thus can be considered to be zero wlog. Therefore our received signal model is

$$y_a(t) = \left(\sum_{n=0}^{N-1} s_n h_a(t - nT_s) \right) e^{2i\pi(f_0 t + \phi_0)} + b_a(t)$$

Once again the likelihood can be written as follows

$$p(y(t)|\phi, f) \propto e^{-\int_{\mathbb{R}} |y_a(t) - \sum_{n=0}^{N-1} s_n h_a(t - nT_s) e^{2i\pi(ft + \phi)}|^2 dt / 2N_0}$$

Recall that the sequence $\{s_n\}$ is known.

So the joint ML estimator for the phase and CFO takes the following form

$$[\hat{\phi}_N, \hat{f}_N] = \arg \min_{\phi, f} \int_{\mathbb{R}} \underbrace{\left| y_a(t) - \sum_{n=0}^{N-1} s_n h_a(t - nT_s) e^{2i\pi(ft + \phi)} \right|^2}_{J_N(\phi, f)} dt$$

By assuming that the CFO is small compared to the bandwidth and by developing the square, we obtain that

$$J_N(\phi, f) = -2\Re \left\{ \sum_{n=0}^{N-1} s_n^* z(n) e^{-2i\pi(fT_s n + \phi)} \right\} + \text{cste} \quad (6)$$

where $z(n) = h_a(-t)^* \star y_a(t)|_{t=nT_s}$ and can be written as follows (under the small CFO assumption)

$$z(n) = s_n e^{2i\pi(f_0 T_s n + \phi_0)} + b(n) \quad (7)$$

where we can prove that $b(n)$ is still a white noise process due to the square-root Nyquist property of the shaping filter.

By letting

$$\alpha_N(f) = \frac{1}{N} \sum_{n=0}^{N-1} s_n^* z(n) e^{-2i\pi f T_s n}$$

we have

$$J_N(\phi, f) = -2N \Re\{\alpha_N(f) e^{-2i\pi\phi}\} + \text{cste}$$

Then, it is easy to check that the term ϕ minimizing $J_N(\phi, f)$ for a given f , is

$$\hat{\phi}_N = \frac{1}{2\pi} \angle(\alpha_N(f)) = \frac{1}{2\pi} \arctan\left(\frac{\Im(\alpha_N(f))}{\Re(\alpha_N(f))}\right) \quad (8)$$

where \angle stands for the phase of a complex-valued number. Now by inserting Eq. (8) into Eq. (6), we can easily show the frequency estimator is obtained by maximizing the modulus of $\alpha_N(f)$. Therefore the joint DA ML phase and CFO estimates are as follow

$$\hat{f}_N = \arg \max_f \left| \frac{1}{N} \sum_{n=0}^{N-1} s_n^* z(n) e^{-2i\pi f T_s n} \right|^2 \quad \text{and} \quad \hat{\phi}_N = \frac{1}{2\pi} \arctan\left(\frac{\Im\left(\frac{1}{N} \sum_{n=0}^{N-1} s_n^* z(n) e^{-2i\pi \hat{f}_N T_s n}\right)}{\Re\left(\frac{1}{N} \sum_{n=0}^{N-1} s_n^* z(n) e^{-2i\pi \hat{f}_N T_s n}\right)}\right). \quad (9)$$

While the phase estimate is in closed-form, the CFO estimate still needs a maximization step. Actually the function to be maximized is the "periodogram". As in the case of timing, this (periodogram) maximization step may be carried out in two steps: the coarse step is done by a FFT of size N . The resulting frequency estimate enables us to initialize a gradient-descent algorithm around the true point, or to use a zoom-FFT.

Remark: The estimators have been developed by developing the MLE based on the continuous-time received signal. Another way is as follows. The received signal can be viewed as

$$y_a(t) = \sum_{n=0}^{N-1} s_n e^{2i\pi(f_0 t + \phi_0)} \Phi_n(t) + b_a(t)$$

where $\Phi_n(t) = h_a(t - nT_s)$. As the shaping filter is a square-root Nyquist filter, the basis functions $\Phi_n(t)$ are orthogonal. In the absence of CFO, the useful part of the received signal is generated by $\Phi_n(t)$; hence, we can project the received signal onto these basis functions without loss of information on the data. Let $u(n) = \langle y_a(t) | \Phi_n(t) \rangle$ with $\langle \cdot | \cdot \rangle$ the inner product. One can check that $u(n) = z(n)$ where $z(n)$ is given in Eq. (7). Developing the ML estimator of the phase and CFO based on $z(n)$ will lead fortunately to the equations reported above. We can also prove that the CRBs obtained by both approaches are identical.

As an illustration, we plot in Fig. 9 (left) a realization of 1000 samples of $z(n)$ when $\phi_0 = 0.1$ and $f_0 T_s = 0.01$ at $E_b/N_0 = 30\text{dB}$ with BSPK modulation. In Fig. 9 (right) the corresponding cost function $J_N(\phi, f)$ has been displayed for $N = 100$. A peak can be observed around the true values of ϕ_0 and f_0 .

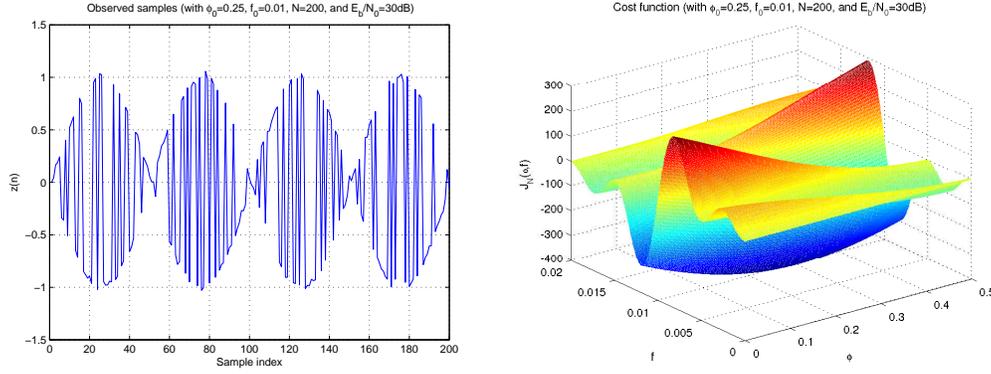


Figure 9: Samples $z(n)$ (left) and corresponding cost function $(\phi, f) \mapsto J_N(\phi, f)$ (right).

Let us now analyze the CRB associated with this estimation issue. In case of estimation of more than one parameter, the CRB is defined through a matrix as follows

$$\text{CRB}(\phi, f) = F^{-1}(\phi, f)$$

where $F(\phi, f)$ is the Fisher Information Matrix whose components here are

$$F(\phi, f) = - \begin{bmatrix} \frac{\partial^2 \log p(y|\phi, f)}{(\partial \phi)^2} & \frac{\partial^2 \log p(y|\phi, f)}{\partial \phi \partial f} \\ \frac{\partial^2 \log p(y|\phi, f)}{\partial f \partial \phi} & \frac{\partial^2 \log p(y|\phi, f)}{(\partial f)^2} \end{bmatrix}$$

After straightforward algebraic manipulations, we obtain the CRB associated with the phase, that is defined as the first element of the diagonal of $\text{CRB}(\phi, f)$, as [32]

$$\text{CRB}(\phi) = \frac{N_0 w_2}{4\pi^2 N (w_0 w_2 - w_1^2)}$$

where, for any integer k

$$w_k = \frac{1}{N^{(k+1)}} \sum_{n=0}^{N-1} n^k |s_n|^2.$$

The CRB associated with the frequency is the second element of the diagonal of $\text{CRB}(\phi, f)$, and is given by

$$\text{CRB}(f) = \frac{N_0 w_0}{T_s^2 4\pi^2 N^3 (w_0 w_2 - w_1^2)}.$$

These expressions can be simplified if $N \gg 1$. We then obtain the so-called asymptotic CRB. One can prove that

$$w_k \xrightarrow{a.s.} \frac{\sigma_s^2}{k+1}.$$

Consequently, we have

$$\text{CRB}(\phi) \approx \frac{1}{\pi^2} \frac{N_0}{E_s} \frac{1}{N}, \text{ and } \text{CRB}(f) \approx \frac{3}{T_s^2} \frac{N_0}{\pi^2} \frac{1}{E_s} \frac{1}{N^3}. \quad (10)$$

Notice that in [33], the (large samples) MSEs of the ML have been calculated and are identical to those given by the asymptotic CRB. Consequently the ML is asymptotically efficient and thus we do not need to spend time to design other estimators since the ML is almost optimal and can be implemented easily in practice.

Thanks to these expressions, some insights about the estimation performance can be given.

- The phase MSE is $\mathcal{O}(1/N)$ and $\mathcal{O}(1/\text{SNR})$. The behavior mimics that of the channel estimator. This is logical since the phase rotation can be viewed as a one-tap linear filter.
- In contrast, the MSE for the CFO decreases much faster as $\mathcal{O}(1/N^3)$. The convergence speed seems to be very high. We will see that we need to have this high convergence speed for the system to operate properly. Let us consider a frame with a training sequence of length N_T followed by a data sequence of length N_D . For the system to operate well, the phase rotation due to the CFO should be kept as low as possible. At the end of the frame, the phase rotation (after correction) is $2\pi T_s(\hat{f}_N - f)(N_T + N_D)$ which is of order $\mathcal{O}((N_T + N_D)/N_T^{3/2})$. Assume a constant ratio β between N_T and N_D . Thus $\beta = N_T/N_D$ and corresponds to the loss in spectral efficiency caused by the training. Then the phase rotation is proportional to $1/\sqrt{N_D}$ and this tends to zero when the frame is large enough. Notice that if the frequency MSE was $\mathcal{O}(1/N^p)$ with $p \leq 2$, the system cannot perform well due to the unbounded phase rotation associated with CFO.

The influence of the noise is similar since the MSE is of order $\mathcal{O}(1/\text{SNR})$.

In Fig. 10, we plot the MSE of the phase and CFO ML estimate and the corresponding CRB: versus E_b/N_0 (left) with $N = 32, T_s = 1\text{s}$, and versus N (right) with $E_b/N_0 = 3\text{dB}$ and $T_s = 1\text{s}$. The training sequence was BPSK modulated.

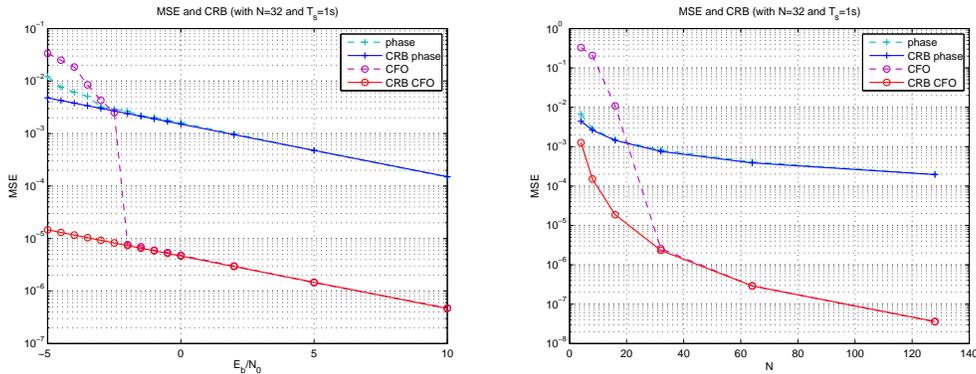


Figure 10: MSE and CRB for phase and CFO versus E_b/N_0 (left) and N (right).

We first observe that when the SNR and the number of samples are high enough, MLE performance perfectly matches the CRB as expected. We also observe a strange phenomenon at low SNR and/or when the number of samples is not large enough: There is a mismatch between the theoretical performance and the empirical one. Moreover the ML

is no longer efficient. This is the so-called outliers effect [34, 35, 36, 37]. It is associated with the failure of the first step of the periodogram maximization. This phenomenon has been analyzed in the literature, and modified expressions for MSE, taking into account this effect have been derived. The most interesting question is: can we fill up the gap between the ML and the CRB by using another estimator? The answer is no. To answer this, some other lower bounds have been developed and analyzed. One can mention the Barankin-like bounds [38, 39, 40, 41, 42], the Bhattacharya-like bounds [43], and the Ziv-Zakai-like bounds [44, 45, 46, 47]. A lot of work has been done on deriving such bounds for the harmonic retrieval issue, where it has been shown that the CRB was not tight at low SNR. Other bounds (especially the Ziv-Zakai one) are actually very close to ML performance, so that it is hard to find better estimates.

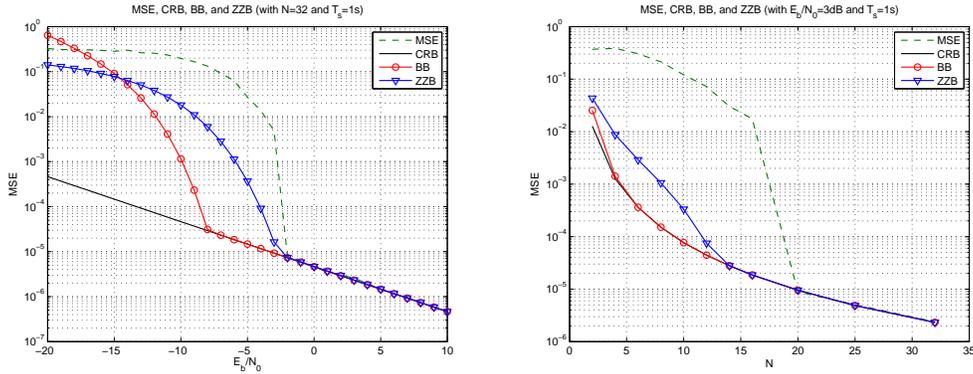


Figure 11: MSE, CRB, BB, and ZZB for CFO versus E_b/N_0 (left) and N (right).

In Fig. 11, we plot the Cramer-Rao Bound (CRB), the (first-order) Barankin Bound (BB), the Ziv-Zakai Bound (ZZB) and the MSE of the ML estimator for CFO parameter versus E_b/N_0 (left) when $N = 32$ and versus N (right) when $E_b/N_0 = 3\text{dB}$. Notice that the threshold (from which the ML performance can be distinguished from the CRB) can be moved to the left by increasing the SNR (when N is fixed) or by increasing N (when the SNR is fixed) in order to obtain the target performance.

Let us now move on to the adaptive version of the ML. As the CFO can be viewed as a phase variation (with a very specific structure), the adaptive algorithm which has the ability of tracking phase variation, has been almost always developed under the assumption that only the phase is non-zero and that the CFO is zero. Therefore we work with the following discrete-time signal:

$$z(n) = s_n e^{2i\pi\phi} + b(n) \quad (11)$$

and the (blockwise) ML for the phase leads to the following cost function (cf. Eq.(6))

$$\hat{\phi}_N = \arg \max_{\phi} \sum_{n=0}^{N-1} \Re \{ s_n^* z(n) e^{-2i\pi\phi} \}$$

Following an approach similar to that for timing, $\hat{\phi}_n$, the estimate of the phase at the n -th

iteration of the (stochastic) gradient algorithm, is updated as follows

$$\begin{aligned}\hat{\phi}_n &= \hat{\phi}_{n-1} + \mu \left. \frac{\partial \Re \{ s_n^* z(n) e^{-2i\pi\phi} \}}{\partial \phi} \right|_{\hat{\phi}_{n-1}} \\ &= \hat{\phi}_{n-1} + \mu \underbrace{\Im \{ s_n^* z(n) e^{-2i\pi\hat{\phi}_{n-1}} \}}_{e(n)}\end{aligned}\quad (12)$$

Thanks to this update equation, we can introduce the famous digital Phase-Locked Loop (PLL) framework.

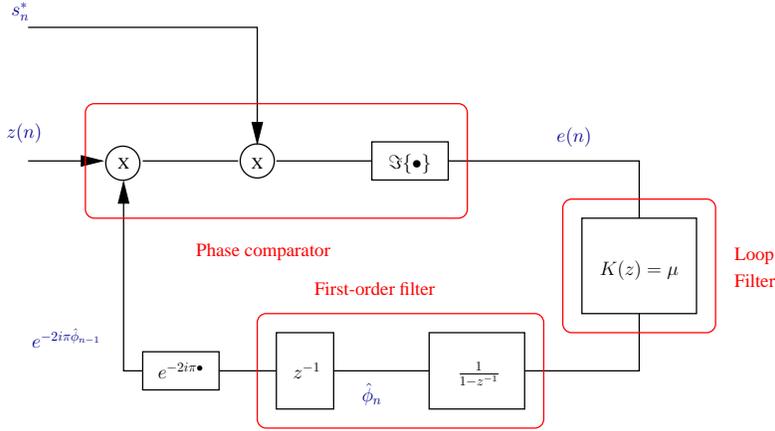


Figure 12: PLL scheme (in DA context).

Note that $\hat{\phi}_n - \hat{\phi}_{n-1}$ can be written as $[1 - z^{-1}] \cdot \hat{\phi}_n$ which means that $\hat{\phi}_n$ is obtained after the phase error $e(n)$ is passed through a filter whose Z-transform is $z \mapsto 1/(1 - z^{-1})$. Therefore the PLL scheme can be described as in Fig. 12.

We would now like to inspect the convergence of the PLL first when the phase is fixed, and then when the phase is time-varying. To do this, we consider that ϕ_0 is fixed for the first 1000 samples $z(n)$, and then is time-varying for the next 1000 samples as follows

$$\phi_{0,n} = \phi_{0,n-1} + w_n, \text{ for } n = 1001, \dots, 2000$$

where $\phi_{0,1000} = \phi_0$ and w_n is a real-valued i.i.d. Gaussian process with zero-mean and variance σ_w^2 . In Fig. 13, we display a realization of the estimated $\hat{\phi}_n$ and the true phase $\phi_{0,n}$ with $\sigma_w^2 = 10^{-4}$ and $E_b/N_0 = 20\text{dB}$.

We remark that proper choice of μ is crucial for good PLL performance. If μ is high, the PLL will rapidly reach an interval around the true value but then will oscillate around the true point without converging. Moreover, a high value of μ enables us to efficiently track the phase variation/noise. In contrast, when μ is small, the convergence speed is low but the phase estimate does not oscillate very much around the true point. But the small value of μ prevents us from following the (too-fast) phase variation/noise.

Indeed, if the phase to be estimated is fixed, it is well known that it is best to consider a time-varying step size μ_n satisfying $\sum_n \mu_n = +\infty$ and $\sum_n \mu_n^2 < +\infty$ [27]. Therefore an

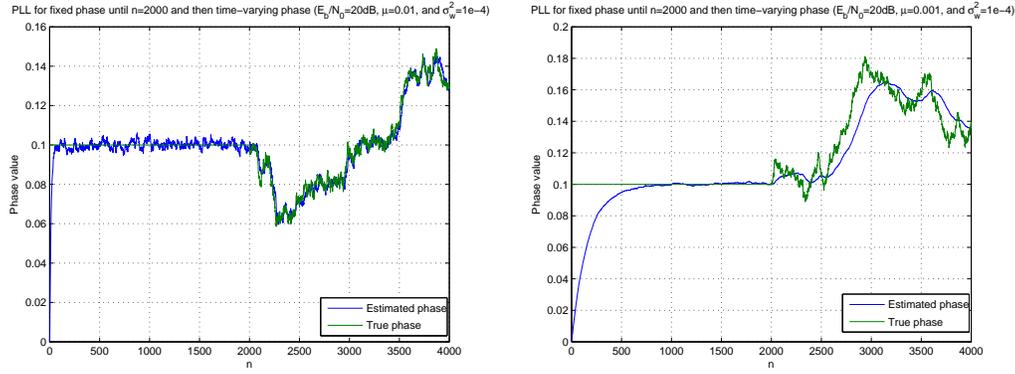


Figure 13: Estimated phase and true phase versus the number of iterations/samples for $\mu = 0.01$ (left) and $\mu = 0.001$ (right).

appropriate choice is $\mu_n = \mu/n$ when the phase is fixed. However such a choice leads to very poor performance if the phase becomes time-varying. A lot of work has been done on design of μ_n adapted to phase noise [27].

2.2. NDA case

Here we will develop NDA estimators

- for timing insensitive to phase and CFO (Problem 2'),
- for phase and frequency insensitive to timing (Problem 2).

As before, we will start with the ML estimators. As in DA mode, the global estimation issue (for timing, phase and frequency) is not tractable. Therefore, we will focus on

- NDA timing ML estimator when phase and CFO are known and thus zero wlog.
- NDA phase and CFO ML estimators when timing is known and thus zero wlog.

NDA ML timing estimator

Now the symbols s_n are unknown. We just assume that s_n belongs to a particular constellation with P states (P -PAM, P -PSK, P -QAM) and that each constellation point is equally likely. We denote by $\{s^{(p)}\}_{p=0, \dots, P-1}$ the set of these constellation points. We also denote by $\mathbf{s}_N = [s_0, \dots, s_{N-1}]^T$ the set of transmitted data symbols. The likelihood takes a much more complicated form due to the need to average over all the potential symbols. Therefore we have

$$p(y|\tau) = \int \cdots \int p(y|\tau, \mathbf{s}_N) p_{\mathbf{s}}(\mathbf{s}_N) d\mathbf{s}_N$$

where $p_{\mathbf{s}}(\cdot)$ is the probability density function of \mathbf{s}_N . Obviously, we have

$$p(y|\tau, \mathbf{s}_N) \propto e^{-\int_{\mathbb{R}} |y(t) - \sum_{n=0}^{N-1} s_n h_a(t - nT_s - \tau)|^2 dt / 2N_0}.$$

and

$$p_{\mathbf{s}}(\mathbf{s}_N) = \prod_{n=0}^{N-1} \left(\frac{1}{P} \sum_{p=0}^{P-1} \delta(s_n - s^{(p)}) \right)$$

as s_n are i.i.d..

After simple but tedious algebraic manipulations (using the fact that the shaping filter is a square-root Nyquist filter), we obtain

$$\hat{\tau}_N = \arg \max_{\tau} \sum_{n=0}^{N-1} \log \left(\frac{1}{P} \sum_{p=0}^{P-1} e^{-\frac{|s^{(p)}|^2}{2N_0}} e^{\frac{\Re\{(s^{(p)})^* z_{\tau}(n)\}}{N_0}} \right)$$

When BPSK is employed, we obtain that

$$\hat{\tau}_N = \arg \max_{\tau} \sum_{n=0}^{N-1} \log \left(\cosh \left(\frac{\Re\{z_{\tau}(n)\}}{N_0} \right) \right)$$

At low SNR, the Taylor series expansion of $\log(\cosh(\cdot))$ can be used to yield

$$\hat{\tau}_N = \arg \max_{\tau} \sum_{n=0}^{N-1} (\Re\{z_{\tau}(n)\}^2 + |z_{\tau}(n)|^2) \quad (\text{BPSK at low SNR}).$$

When QPSK is employed, similar derivations lead to

$$\hat{\tau}_N = \arg \max_{\tau} \sum_{n=0}^{N-1} |z_{\tau}(n)|^2 \quad (\text{QPSK at low SNR}).$$

We will see that the NDA ML approach for timing is very complicated except for BPSK and QPSK. To handle constellations with more states, Expectation-Maximization (EM) algorithm can be employed and will be explained later in the section devoted to Code-aided synchronization. Notice that in [48] EM algorithm is also implemented but in another way: indeed, the timing error is viewed as the nuisance parameter with an *a priori* distribution while the data are viewed as the useful parameters to be detected. The authors thus attempt to extend the Maximum Likelihood Sequence Estimator to the case of timing error and are able to correct it through an iterative implementation.

We observe, from Fig. 14, that this timing estimator is very sensitive to phase and CFO. Therefore, we do not continue its analysis in depth. Obviously, an adaptive version may be implemented and the early-late trick can be employed here.

It is clear that in the NDA case, we need to develop a sub-optimal (non-ML) timing estimator which is insensitive to phase and CFO. We do so next.

NDA ML phase and CFO estimator

We now assume that the timing is known and thus zero wlog. Once again, we would like to characterize the NDA joint ML phase and CFO estimator.

By following the same reasoning as for the timing, we obtain that

$$[\hat{\phi}_N, \hat{f}_N] = \arg \max_{\phi, f} \sum_{n=0}^{N-1} \log \left(\frac{1}{P} \sum_{p=0}^{P-1} e^{-\frac{|s^{(p)}|^2}{2N_0}} e^{\frac{\Re\{(s^{(p)})^* z(n) e^{-2i\pi(fT_s n + \phi)\}}}{N_0}} \right). \quad (13)$$

Figure 14: MSE for NDA ML based timing estimator for BPSK versus phase mismatch (left) and CFO mismatch (right).

In contrast with the DA case, we are not able to write $\hat{\phi}$ with respect to f . As a consequence, the maximization remains a 2-D search which is extremely time consuming. Therefore, we could focus on i) CFO estimation issue when phase is known, and ii) phase estimation when CFO is known. The first scheme is clearly unrealistic, and thus will be omitted. In contrast, the second scheme is of interest, especially in the tracking regime.

We now assume that **CFO is known and thus can be assumed to be zero wlog.** Eq. (13) can be simplified as follows

$$\hat{\phi}_N = \arg \max_{\phi} \sum_{n=0}^{N-1} \log \left(\frac{1}{P} \sum_{p=0}^{P-1} e^{-\frac{|s^{(p)}|^2}{2N_0}} e^{\frac{\Re\{(s^{(p)})^* z(n) e^{-2i\pi\phi}\}}{N_0}} \right). \quad (14)$$

In Fig. 15, we have plotted the cost function given in Eq. (14) for various QAM constellations. Notice that the smaller the constellation, the sharper is the cost function. Therefore it is easier to estimate phase for small constellation sizes. When a high-order constellation is used, the SNR must be high enough to ensure accurate synchronization (this is not necessarily a drawback since high-order modulation requires high SNR for detection) and the number of samples must be large enough as well (which clearly may become an issue). Fig. 16 depicts the, MSE of NDA ML based phase estimate versus SNR, N and the timing error respectively for BPSK constellation. The MSE decreases proportionally to $1/\text{SNR}$ and $1/N$. Moreover it is insensitive to timing error.

To overcome the implementation issues due to the highly complicated shape of the cost function, EM algorithm can be employed as in [49].

Once again, as with the NDA ML based timing estimate, the cost function in Eq. (14) can be simplified when BPSK is used.

$$\hat{\phi}_N = \arg \max_{\phi} \sum_{n=0}^{N-1} \log \left(\cosh \left(\frac{\Re\{z(n) e^{-2i\pi\phi}\}}{N_0} \right) \right) \quad (\text{BPSK}) \quad (15)$$

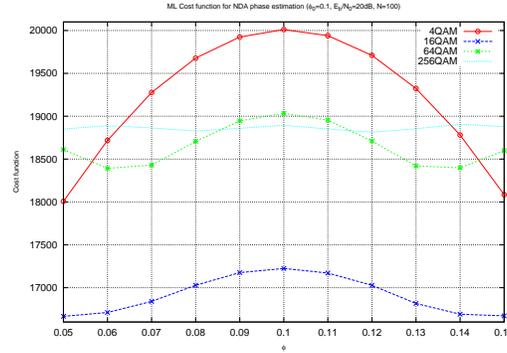


Figure 15: Cost function of NDA ML phase estimate (with $\phi_0 = 0.1$, $E_b/N_0 = 20\text{dB}$, $N = 100$) for various P -QAM.

An exhaustive search to solve Eq. (15) would be quite time consuming. However, at low SNR, as with the timing case, we obtain that

$$\hat{\phi}_N = \arg \max_{\phi} \sum_{n=0}^{N-1} \Re\{z(n)^2 e^{-2i\pi 2\phi}\} \Leftrightarrow \hat{\phi}_N = \frac{1}{2} \angle \left(\sum_{n=0}^{N-1} z(n)^2 \right) \quad [\pi] \quad (\text{BPSK at low SNR}).$$

The NDA ML phase estimate is equivalent to the so-called Square-Power estimate at low SNR which can be easily implemented. Similar results can be obtained for any P -QAM or any P -PSK. Indeed, the NDA ML phase estimate is equivalent to the M -th Power estimate with $M = 4$ for any P -QAM and $M = P$ for any P -PSK [50]. Further information about other phase estimators, with PSK modulation, can be found in [51].

Let us consider **BPSK** case in more detail. We would like to implement Eq. (15) in an adaptive manner. Let $\hat{\phi}_n$ be the value of the estimated phase at time n (*i.e.*, at the n -th iteration). the update equation is

$$\begin{aligned} \hat{\phi}_n &= \hat{\phi}_{n-1} + \mu \left. \frac{\partial \log \left(\cosh \left(\frac{\Re\{z(n)e^{-2i\pi\phi}\}}{N_0} \right) \right)}{\partial \phi} \right|_{\hat{\phi}_{n-1}} \\ &= \hat{\phi}_{n-1} + \mu \Im \left\{ \check{s}_n z(n) e^{-2i\pi\hat{\phi}_{n-1}} \right\} \end{aligned}$$

with

$$\check{s}_n = \tanh \left(\frac{\Re\{z(n)e^{-2i\pi\hat{\phi}_{n-1}}\}}{N_0} \right). \quad (16)$$

Notice that the update equation is very similar to that in the DA case (cf. Eq.(12)) except that s_n has been replaced with \check{s}_n . Be looking at carefully \check{s}_n , we observe that it corresponds to the so-called soft decision under BPSK constellation assumption. If the soft decision is replaced with a hard decision, then we are in the Decision Directed context, and the estimator \check{s}_n becomes

$$\check{s}_n = \text{sign} \left(\frac{\Re\{z(n)e^{-2i\pi\hat{\phi}_{n-1}}\}}{N_0} \right) \quad (\text{Decision-Directed}).$$

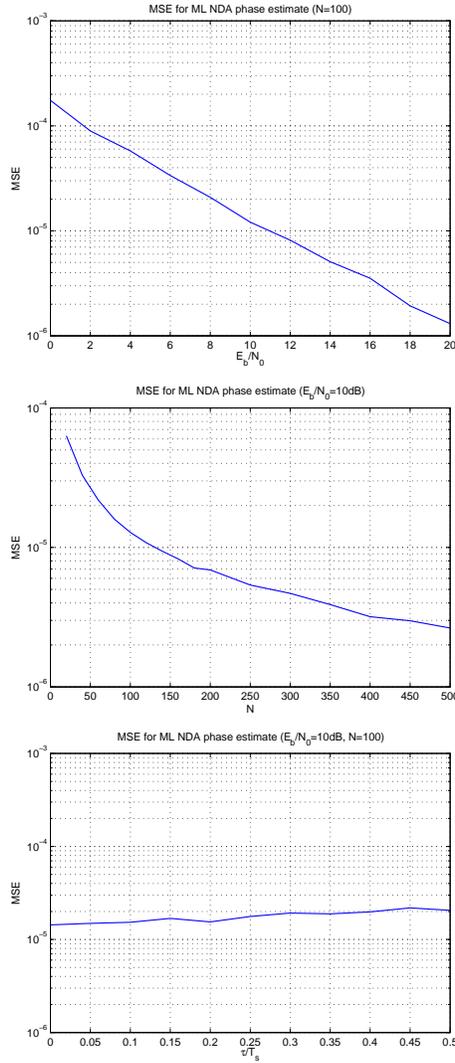


Figure 16: MSE of NDA ML phase estimate with BPSK modulation versus E_b/N_0 (top), N (middle), and τ_0 (bottom).

If a decision is made (other than soft or hard), the term \check{s}_n can be written as

$$\check{s}_n = \frac{\Re\{z(n)e^{-2i\pi\hat{\phi}_{n-1}}\}}{N_0} \quad (\text{No data decision})$$

and corresponds to the famous Costas loop originally introduced for continuous-time amplitude-modulated signals [52, 53].

In conclusion, we recall that the joint NDA sync parameters estimation issue is intractable. To simplify it, we first focus on the timing. The NDA ML timing estimator is quite complicated and very sensitive to phase and CFO. So it is useless in practice. We next consider phase and CFO estimation issues. Clearly the joint problem is still intractable.

Therefore we only focus on the phase estimation issue assuming CFO is known. When the constellation size is small enough, the NDA ML phase is implementable in blockwise or adaptive manner and is not too sensitive to timing error.

Consequently, in order to build practical system, we still need sub-optimal (non-ML) blind methods

- for estimating the timing without the knowledge of the phase and the CFO.
- for estimating the CFO without the knowledge of the timing and the phase
- for estimating the phase when the constellation size becomes high. The insensitivity to the timing and the CFO is not required since the phase estimator is often carried out after timing and CFO correction.

Sub-optimal estimators

In this section, we will develop sub-optimal estimators for the sync parameters. To do that, we will analyze carefully some statistical properties of the received signal. Based on this properties, we will be able to introduce some powerful estimators. The main statistical tools we need are briefly described below

- **The Cyclostationarity tool:** for sake of simplicity, we will only define second-order cyclostationarity. For more details about any-order cyclostationarity, the reader may refer to [54, 55]. Let us consider $x(t)$ a (complex-valued) continuous-time process and its correlation function $(t, \tau) \mapsto r(t, \tau) = \mathbb{E}[x(t + \tau)x(t)^*]$. When the process is second-order stationary, then the function $(t, \tau) \mapsto r(t, \tau)$ is independent of t . In contrast, if the function $(t, \tau) \mapsto r(t, \tau)$ is periodic with respect to t (and the period is independent of τ), the process is said to be cyclostationary. Let T_c be the period of the correlation function. T_c is called the cyclic period. By doing a Fourier series expansion, we have

$$r(t, \tau) = \sum_k r^{(k)}(\tau) e^{2i\pi k t / T_c}$$

where

- $f_k = k/T_c$ is the so-called k -th cyclic frequency,
- $\tau \mapsto r^{(k)}(\tau)$ is the cyclic correlation (at cyclic frequency k),
- $S^{(k)}(e^{2i\pi f}) = \sum_{\tau} r^{(k)}(\tau) e^{-2i\pi f \tau}$ is the cyclic spectrum (at cyclic frequency k).

A similar analysis can be done with the pseudo-correlation $(t, \tau) \mapsto u(t, \tau) = \mathbb{E}[x(t + \tau)x(t)]$ if it is not identically zero.

- **The Non-Circularity tool:** let us consider a random (complex-valued) zero-mean scalar variable x . A variable is said to be circularly-symmetric (or simpler circular) if x and $\tilde{x} = e^{2i\pi\theta}x$ have the same distribution regardless of the rotation angle θ . Consequently the distribution is rotationally invariant. By assuming the moment at any order exists, we have $\mathbb{E}[x^p x^{*q}] = \mathbb{E}[\tilde{x}^p \tilde{x}^{*q}]$ which implies that $\mathbb{E}[x^p x^{*q}] = 0$ as soon as $p \neq q$. Thus, when a variable is circularly-symmetric, only the moment at even

orders may be non-zero. Moreover only the moments depending on a power of $|x|$ are non-zero. In digital communications, many variables are circular not for arbitrary orders but only until a certain order. If

$$\mathbb{E}[x^p x^{*q}] = 0$$

for $p \neq q$ and $p+q < M$, the random variable x is said to be circular until the $(M-1)$ -th order or equivalently to be non-circular from the M -th order. For more details about the non-circularity, the reader may refer to [56, 50, 12].

The rest of this section is now organized as follows: we focus on the timing estimator (based on the cyclostationary tool), then the CFO estimator (based on the cyclostationary or non-circularity tools), and the phase estimator (based on the non-circularity tool).

Timing estimation issue:

We observe that the signal $y_a(t)$ is cyclostationary with period T_s . Consequently the discrete-time (over-sampled) signal $\tilde{y}(n) = y_a(nT_s/Q)$ is also cyclostationary with period Q , *i.e.*, the sequence $\mathbb{E}[\tilde{y}(n+m)\tilde{y}(n)^*] = \mathbb{E}[\tilde{y}(Q+n+m)\tilde{y}(Q+n)^*]$ is period for any integer m . Notice that if the continuous-time received signal is sampled at the symbol rate, $Q = 1$ and thus $y(n)$ is stationary. In contrast, as soon as $Q \geq 2$, $\tilde{y}(n)$ is strictly cyclostationary.

First of all assuming that the CFO is zero, it has been remarked in [57] that $r^{(k)}(m) = \alpha_{k,m} e^{-2i\pi k\tau/T_s}$ for $k \neq 0$, with $\alpha_{k,m}$ a known complex-valued constant depending on the shaping filter. Thanks to this relationship between $r^{(k)}(m)$ and τ , the following ad hoc estimator was proposed in [57]:

$$\hat{t}_N = -T_s \sum_{k \in \mathbb{K}^*} \sum_{m \in \mathbb{M}} \frac{1}{k} \angle \left(\hat{r}^{(k)}(m) \alpha_{k,m}^{-1} \right) \quad (17)$$

where \mathbb{K}^* is any set of integers, not including 0, and \mathbb{M} is any set of integers m such that $r^{(k)}(m) \neq 0$. The term $\hat{r}^{(k)}(m)$ is the empirical estimate of $r^{(k)}(m)$ based on N samples, and is given by

$$\hat{r}^{(k)}(m) = \frac{1}{N} \sum_{n=0}^{N-1} \tilde{y}(n+m)\tilde{y}(n)^* e^{-2i\pi nk/Q} \quad (18)$$

The performance of such an empirical cyclic correlation estimate can be found in [16, 58, 59, 57].

Notice that the timing estimator is insensitive to the phase, and thus can be used before the phase estimation step. In contrast, when the CFO is not zero, it must be estimated first; the timing estimate is then modified as follows

$$\hat{t}_N = -T_s \sum_{k \in \mathbb{K}^*} \sum_{m \in \mathbb{M}} \frac{1}{k} \angle \left(\hat{r}^{(k)}(m) \alpha_{k,m}^{-1} e^{-2i\pi m \hat{f}_N T_s / Q} \right) \quad (19)$$

where \hat{f}_N is the CFO estimate.

When $\mathbb{K} = \{1\}$, $\mathbb{M} = \{0\}$, and the conjugate operation on \tilde{y} is removed, we obtain the heuristic algorithm introduced in [60] which actually is well-suited for BPSK constellation. When $\mathbb{K} = \{1\}$ and \mathbb{M} is any singleton, we obtain the algorithm in [61]. With $\mathbb{K} = \{-1, 1\}$

and $\mathbb{M} = \{0\}$, the estimator reduces to that in [62]. This last algorithm has the great advantage of being insensitive to the CFO since, for $m = 0$, Eq. (17) is identical to Eq. (19).

All these timing estimators have been theoretically analyzed and closed-form expressions for the Mean-Square Error are given in [63].

The question now is: what is the best blind estimator for timing relying on the cyclic correlation. When $P = 2$, the answer can be found in [63] and corresponds to the weighted covariance matching principle [64, 65]. The gain in performance is only incremental at the expense of higher complexity. Therefore the NDA timing estimator proposed in Eq. (17) (especially that in [62] which is insensitive to CFO) is a strong candidate for our problem.

CFO estimation issue:

The cyclostationarity property can also be used for estimating the CFO blindly. Indeed, in [57], it has been remarked that $r^{(k)}(m) = \alpha_{k,m} e^{2i\pi k\tau/T_s} e^{2i\pi m f T_s/Q}$ with β_k a complex valued scalar which induces the following estimator

$$\hat{f}_N = \frac{P}{4\pi T_s} \sum_{k \in \mathbb{K}^*} \sum_{m \in \mathbb{M}^*} \frac{1}{m} \angle \left(\hat{r}^{(k)}(m) \hat{r}^{(-k)}(m) \alpha_{k,m}^{-1} \alpha_{-k,m}^{-1} \right) \quad (20)$$

where \mathbb{M}^* corresponds to any set of integers that does not contain zero. This estimator is insensitive to timing and phase. Once again, by taking specific value of the set \mathbb{K} and \mathbb{M} , we obtain the estimators introduced in [61] and [62]. The analytical analysis of the estimator has been done in [66].

Another approach to estimate the CFO independent of the timing and the phase is to use the non-circularity property of the received signal $y_a(t)$. Indeed, it is easy to check [50] that

$$\mathbb{E}[s_n^M] \neq 0$$

with

- $M = 2$ for P -PAM constellation
- $M = P$ for P -PSK constellation
- $M = 4$ for P -QAM constellation

Consequently the receive signal is also non-circular at M -th order. For the sake of simplicity, let us consider that the receive signal has been sampled at the symbol rate after passing through the matched filter. Extension to the oversampled case is straightforward. Then

$$y(n) = \left(\sum_{\ell=0}^L g(\ell) s_{n-\ell} \right) e^{2i\pi f T_s n} + b(n)$$

where $g(n) = \tilde{h}_a(nT_s - \tau_0) e^{2i\pi\phi_0}$, $(L+1)$ is the channel length and $b(n)$ the additive white Gaussian noise. Notice that if the timing has been perfectly corrected and the shaping filter is a square-root Nyquist filter, we have $g(n) = \delta_{0,n} e^{2i\pi\phi_0}$ and thus $L = 0$. When the timing is not fully corrected, the filter $g(n)$ is no longer a one-tap filter and ISI occurs. Moreover as τ_0 and ϕ_0 are unknown, we have to assume that $g(n)$ is unknown as well.

We can rewrite $y(n)$ as

$$y(n) = a(n)e^{2i\pi f T_s n} + b(n) \quad \text{where} \quad a(n) = \sum_{\ell=0}^L g(\ell)s_{n-\ell}$$

Our estimation problem is then equivalent to the estimation of a harmonic f disturbed by additive white Gaussian noise $b(n)$ and multiplicative noise $a(n)$. A lot of work has been done on the problem of harmonic retrieval in multiplicative and additive noise. We will summarize the essential results.

Let us first assume P -PAM constellation. Let $u_a(m) = \mathbb{E}[a(n+m)a(n)]$ be the pseudo-correlation at lag m . As PAM is employed, there exists at least one lag such that $u_a(m)$ is non-zero. Let $p_m(n) = y(n+m)y(n)$. We have

$$p_m(n) = u_a(m)e^{2i\pi(2fT_s)n} + e_m(n)$$

where $e_m(n)$ is a zero-mean process which can be viewed as noise. But this noise is not Gaussian, nor white, nor stationary. Nevertheless by working with $p_m(n)$, we now have to estimate an harmonic ($2f$) disturbed only by additive noise. The multiplicative noise has been removed. We recall that if $e(n)$ is Gaussian white and stationary, the ML based on $p_m(n)$ will lead to peak-picking the periodogram. Even if $e(n)$ does not satisfy these standard assumptions, it is still usual practice to estimate the frequency through peak-picking the periodogram (even if it no longer has any link with the ML). Then, we have

$$\hat{f}_N = \arg \max_f \sum_m \left| \frac{1}{N} \sum_{n=0}^{N-1} p_m(n) e^{-2i\pi(2f)T_s n} \right|^2.$$

When $m = 0$, we have

$$\hat{f}_N = \arg \max_f \left| \frac{1}{N} \sum_{n=0}^{N-1} y(n)^2 e^{-2i\pi(2f)T_s n} \right|^2$$

which is the well-known (and classical) square-power estimator.

To analyze the theoretical performance of these estimators, the standard work on harmonic retrieval (which assumes the additive noise is Gaussian) cannot be applied [67, 68, 69]. However the analysis has been done quite recently in [70]. Notice that the above approach holds if $a(n)$ is real-valued without assuming any specific structure [59, 71]. However, the previous estimator has limited impact since the PAM constellation is not spectrally efficient.

We now consider P -PSK and P -QAM constellations. Let M be the non-circularity order. Recall that $M = P$ for PSK and $M = 4$ for QAM. Let $q(n) = y(n)^M$. Then

$$q(n) = \mathbb{E}[y^M(n)]e^{2i\pi(Mf)n} + e'(n).$$

where $e'(n)$ is a zero-mean process that can be interpreted as additive noise. Once again one can carry out periodogram peak-picking for retrieving f . Thus, we have

$$\hat{f}_N = \arg \max_f \left| \frac{1}{N} \sum_{n=0}^{N-1} y(n)^M e^{2i\pi(Mf)T_s n} \right|^2. \quad (21)$$

This estimator has been introduced by [72] for PSK. Several extensions of these estimators may be introduced. For example, instead of relying only on $y(n)^M$, one could also work with $y(n)y(n+m_1)\cdots y(n+m_{M-1})$. This has not been done in the literature. One could also find the best non-linear transformation \mathcal{F} such that peak-picking the periodogram based on $\mathcal{F}(y(n))$ has some desired properties. One can find some results about optimization of \mathcal{F} in [73].

In Figs. 17-18, we plot the theoretical and empirical performance of the M -th power estimate for various QAM constellation, varying the SNR and the number of samples N respectively.

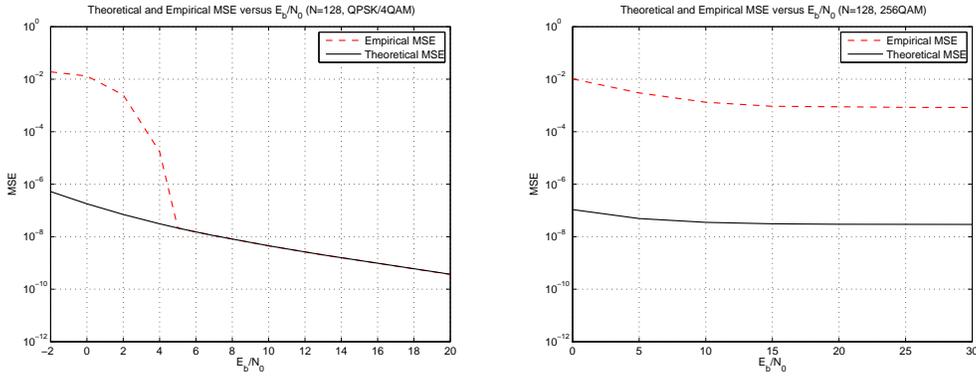


Figure 17: Theoretical and Empirical MSE of M -th power CFO estimate versus E_b/N_0 for 4-QAM (left) and 256QAM (right) with $N = 128$.

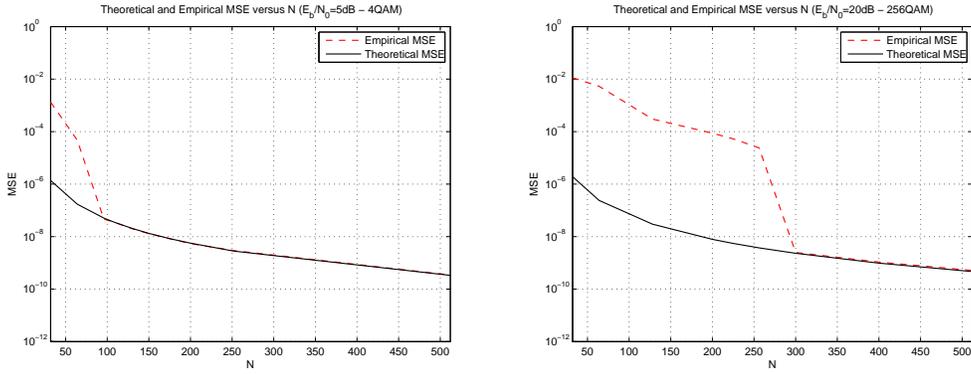


Figure 18: Theoretical and Empirical MSE of M -th power CFO estimate versus N for 4-QAM at $E_b/N_0 = 5\text{dB}$ (left) and 256QAM at $E_b/N_0 = 20\text{dB}$ (right).

For QAM, we observe a self-noise phenomenon since the performance has an error floor with respect to SNR. It is due to the fact that $s_n^4 \neq \mathbb{E}[s_n^4]$. A recent approach can be used to remove the self-noise (see [74]).

The outliers effect still occurs as in any approach based on peak-picking periodogram. The theoretical MSE taking into account the outliers has been analyzed in [37] for modulated signals and [34] for unmodulated signals.

Phase estimation issue:

For estimating the phase, we will assume that the timing and CFO have already been correctly compensated. Then we focus on the sampled (at the symbol rate) output of the matched filter. Therefore, we have

$$z(n) = s_n e^{2i\pi\phi_0} + b(n).$$

The approach for estimating the phase will be similar to that for the estimation of CFO by using the non-circularity property of the constellation. Since

$$q(n) = y(n)^M = \mathbb{E}[s_n^M] e^{2i\pi(M\phi_0)} + e(n)$$

where $e(n)$ is a zero-mean process and $\mathbb{E}[s_n^M]$ is assumed to be known with phase ϕ_* . Then it is easy to build a simple estimator as follows

$$\hat{\phi}_N = \frac{1}{M} \angle \left(\frac{1}{N} \sum_{n=0}^{N-1} y(n)^M \right) - \phi_* \quad (22)$$

Obviously we have an ambiguity of $2\pi/M$ since the constellation is invariant to rotations of angle $2\pi k/M$, for any integer k . We recall that this estimate is close to the ML at low SNR.

This approach was introduced for PSK by [72] and for QAM by [50]. Notice that other less powerful estimates of the phase have been introduced in the literature. A deep theoretical analysis of the estimators can be found in [75].

CRB

In the DA case, we have observed that the derivations of the CRB was not a difficult task. Moreover, we were able to obtain rather nice closed-form expression that enable us to provide some interesting insights. In the NDA case, in contrast, the CRB derivations in closed-form is much more difficult due to the presence of the unknown data sequence. We thus work in the framework of the CRB with nuisance parameters. Obviously the nuisance parameters here are the data. Why is it difficult? To calculate the CRB, we need at least to have a closed-form expression for the likelihood. As noted earlier in the discussion of blind ML estimators, a closed-form expression of the likelihood is very difficult to obtain.

The main idea is to define a variant of the CRB assuming a specific property for the data sequence (the nuisance parameters). In order to be useful, these CRB variations should be easier to derive and should provide some insights. Typically, these CRB variants will not be as tight as the classical CRB, and hence they will be optimistic.

A lot of work has been done on the derivation of the CRB for synchronization parameters and/or channel estimators [76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86]. We will summarize the main principle in deriving CRB with nuisance parameters. Applications to phase and CFO estimation are provided with more details.

Let us start with the definition of several CRBs :

- **True-TCRB** (also called Unconditional-UCRB or Stochastic-SCRB) : the data sequence is viewed as a random process with a known distribution (a product of sum of Dirac distributions). Let θ be the synchronization parameter vector [82, 83, 84, 86]. The true Fisher Information Matrix is defined as

$$F_t(\theta) = -\mathbb{E} \left[\frac{\partial^2 \log(p(y|\theta))}{\partial \theta (\partial \theta)^T} \right]$$

with the true likelihood

$$p(y, \theta) = \mathbb{E}_{\mathbf{s}} [p(y|\theta, \mathbf{s})] = \int \cdots \int p(y, |\theta, \mathbf{s}) p(\mathbf{s}) d\mathbf{s}$$

Then the true CRB is obtained as

$$\text{TCRB}(\theta) = F_t(\theta)^{-1}$$

The main drawback is that it is usually intractable to express this in closed-form. Some expressions are given in [83] but it is not really in closed-form due to the huge number of sums and products.

- **Conditional-CCRB** (also called Deterministic-DCRB) : the data sequence is assumed to be of interest and are added to the parameters to be estimated. Notice that the structure of the data sequence (i.e. data belong to a specific discrete constellation) is not taken into account. If we would like to take this structure into account, we should calculate the CRB when strong constraints on the parameters have to be satisfied. Such a problem is very hard [87, 88] and thus we are far away from our objective of simplifying the CRB derivations. The difficulty can be partially overcome by considering *Bayesian* CRBs which use *a priori* distributions on the data [24]. Let us come back to the CCRB [76].

$$F_c(\theta) = -\mathbb{E} \left[\frac{\partial^2 \log(p(y|\theta, \hat{\mathbf{s}}_N))}{\partial \theta (\partial \theta)^T} \right]$$

where $\hat{\mathbf{s}}_N$ is obtained by maximizing over \mathbf{s} (by omitting all the constraints on \mathbf{s}) the conditioned likelihood $\mathbf{s} \mapsto p(y|\theta, \mathbf{s})$. Then the conditioned CRB is obtained as

$$\text{CCRB}(\theta) = F_c(\theta)^{-1}$$

Usually the asymptotic version is used i.e., the number of samples is assumed to be large.

- **Gaussian-GCRB**: The true CRB is very difficult to compute due to the distribution of \mathbf{s} . To overcome this problem, one can assume that \mathbf{s} is Gaussian distributed (even if though this is not true). The GCRB is thus obtained by using the same definition of the true CRB but by assuming a Gaussian data sequence. Obviously this GCRB is not generally a bound since it implicitly assumes that the higher-order cumulants are zero. For example, if the data is PSK and QAM, we have seen that the higher-order moments are very important through the M -th power estimate and thus the GCRB

does not capture this information. Nevertheless the GCRB has two advantages : it is quite easy to derive in closed-form, and it is a bound on the MSE for all the estimates based on second-order statistics (even if the data sequence is not Gaussian) [78, 79, 85].

- **Modified-MCRB** : data are assumed to be known in order to calculate the Fisher Information Matrix. In order to have a Fisher Information Matrix (FIM) independent of data, the data-dependent FIM is averaged over the data. Therefore, we have

$$F_m(\theta) = -\mathbb{E}_{\mathbf{s}} \left[\mathbb{E}_y \left[\frac{\partial^2 \log(p(y|\theta, \mathbf{s}))}{\partial \theta (\partial \theta)^T} \right] \right].$$

Then the MCRB [77, 80, 81] is

$$\text{MCRB}(\theta) = F_m(\theta)^{-1}$$

In terms of derivations, the main difference from the TCRB is that the average over the data is outside the log (for MCRB) and inside the log for the TCRB. This quite small difference leads to tractable MCRB and usually intractable TCRB. Indeed, the log will be removed by the exp in $p(y)$ due to the Gaussian additive noise. In MCRB, log will directly remove the exp although it is not the case for TCRB, thus leading to very hard calculations.

In the sequel, we provide some general results on the links between the various CRBs. This relationship is independent of the estimation problem.

- At low SNR, the TCRB is much more tractable by doing a Taylor series expansion of e^x for small x [82, 84, 86].
- At any SNR, the MCRB is a bound but sometimes too optimistic and thus not tight enough.
- At any SNR, we have $\text{TCRB} \geq \text{MCRB}$ and $\text{CCRB} \geq \text{MCRB}$ [76].
- At high SNR, if the data sequence belongs to a discrete set, we have $\text{TCRB}/\text{MCRB} \rightarrow 1$. So in our context where data belong to a finite constellation, this property holds and thus the MCRB is very useful as soon as the SNR is large enough [82].

Let us now focus on our specific estimation problem: for timing estimation (when phase and CFO are perfectly corrected), there already exists a very good tutorial [76]. Therefore we focus on phase and CFO estimation (when timing is known and thus zero wlog). Then we work with the (symbol-rate) sampled output of the matched filter. Thus our signal model is

$$y(n) = s_n e^{2i\pi(f_0 T_s n + \phi_0)} + b(n).$$

Let us focus on a "toy" example to highlight some interesting behaviors. We first assume that s_n can be decomposed as follows

$$s_n = \sigma_R s_n^{(R)} + i \sigma_I s_n^{(I)}$$

where $s_n^{(R)}$ and $s_n^{(I)}$ are two unit-variance white BPSK sequences independent of each other; $\sigma_R^2 = (1+u)/2$ and $\sigma_I^2 = (1-u)/2$ with $u \in [0, 1]$ such that the correlation $\mathbb{E}[|s_n|^2] = 1$ and $\mathbb{E}[s_n^2] = u$. Notice that if $u = 1$, then s_n becomes a standard BPSK. And if $u = 0$, s_n is not a (second-order) non-circular process anymore. Even though both phase and CFO have to be estimated, we only provide the expressions for CRB associated with the CFO estimate.

- **TCRB derivations:** As the nuisance parameter is discrete, we know that at high SNR the TCRB is equivalent to the MCRB; the latter can be calculated very easily by averaging the FIM given in Eq. (10). Therefore, we have

$$\text{TCRB}(f)|_{\text{high SNR}} = \frac{3N_0}{\pi^2 T_s^2 N^3}$$

At low SNR [82], a derivation based on the Taylor series expansion of $x \mapsto e^x$ leads to

$$\text{TCRB}(f)|_{\text{low SNR}} = \frac{3N_0^2}{\pi^2 u^2 T_s^2 N^3}$$

We observe that the non-circularity power (induced by u) has a great impact at low SNR where the performance is proportional to SNR^2/u^2 .

- **CCRB derivations:** as we deal with the (symbol rate) sampled signal $y(n)$, the CCRB corresponds to an underdetermined estimation problem: we have N samples to estimate $N + 2$ parameters, and no additional constraints. As a consequence, the CCRB is not finite. To overcome the problem, we need to work with the oversampled version of the received signal. This is out of the scope of our toy example. For more details, please refer to [89, 76].
- **GCRB derivations:** the expressions can be found in [85] and are reported below

$$\text{GCRB}(f) = \frac{3(1 - u^2 + 4N_0 + 4N_0^2)}{4\pi^2 u^2 T_s^2 N^3} \text{ (for } N \text{ large enough)}$$

Once again the non-circularity power (u) has a great impact which verifies that the non-circularity property is an important tool in blind estimation of frequency and phase. We can even show that when $u \neq 0$ (and especially for $u = 1$, *i.e.*, BPSK), the square-power estimate MSE is identical to the GCRB. Therefore the square-power estimator for non-circular white multiplicative noise is the best second order estimator.

If $u = 0$, s_n is not (second-order) non-circular anymore. The second order statistics reduce to the correlation $r_y(m) = r_s(m)e^{2i\pi f_0 T_s m} + 2N_0\delta(m)$. If s_n is white, we do not have information about f and the GCRB will go to infinity. In contrast, if s_n is colored, we can easily build an estimate as follows: $\hat{f}_N = (\angle \hat{r}_y(m) - \angle(r_s(m)))/m$ where $\hat{r}_y(m)$ is the empirical estimate of $r_y(m)$ and $r_s(m)$ is, known. The frequency is thus viewed as the phase of the correlation function. The GCRB is given by [79]

$$\text{GCRB}(f)|_{\text{circular case}} = \frac{1}{N \int_0^1 \left(\frac{S'_s(e^{2i\pi f})}{S_s(e^{2i\pi f}) + 2N_0} \right)^2 df} \text{ (for } N \text{ large enough)}$$

where $S_s(e^{2i\pi f})$ is the power spectrum of s_n and $S'_s(e^{2i\pi f})$ its derivative function. We note that the GCRB decreases as $1/N$ which is the convergence speed associated with the phase estimation issue. So clearly, to have faster convergence, we need to use high-order statistics of s_n when it is second-order circular.

- **MCRB derivations:** we just have to average the FIM and it is the term given in Eq. (10).

$$\text{MCRB}(f) = \frac{3N_0}{\pi^2 T_s^3 N^3}.$$

The MCRB does not capture the influence of the non-circularity; it is not tight enough and is too optimistic except at high SNR.

Let us now assume that s_n belongs to an iid sequence drawn from a PSK or QAM constellation. The GCRB, MCRB are not modified at all although the process is no longer (second-order) non-circular. In contrast the TCRB (at low SNR) is completely different and is given [82] by

$$\text{TCRB}|_{\text{low SNR, } M\text{-th order non. circ}} = \mathcal{O}(1/\text{SNR}^M)$$

Once again the non-circularity tool is fundamental for estimating the CFO and the phase. The TCRB has similar behavior as the MSE of the M -th power (if outliers effect is not taken into account) at low SNR.

In Figs. 19-20-21, we plot various CRBs and empirical MSE of the M -th power estimate versus SNR, N and u respectively for the process s_n used in the toy example.

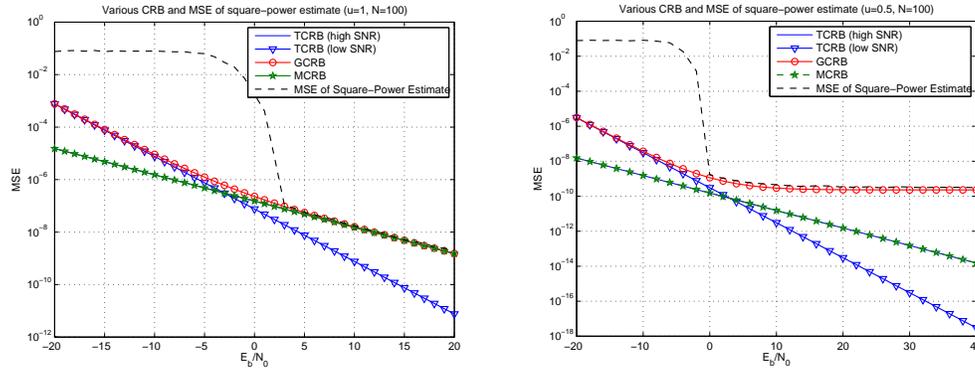


Figure 19: Various CRBs and MSE versus E_b/N_0 for $u = 1$ (left) and for $u = 0.5$ (right).

Code-Aided synchronization

When the SNR is low, the NDA estimators may offer poor performance with realistic number of samples. Until now, we have not exploited the usual structure of the data. Indeed, in order to obtain the targeted BER in current systems, channel coding is used. There are two ways of using the channel structure to improve the synchronization step

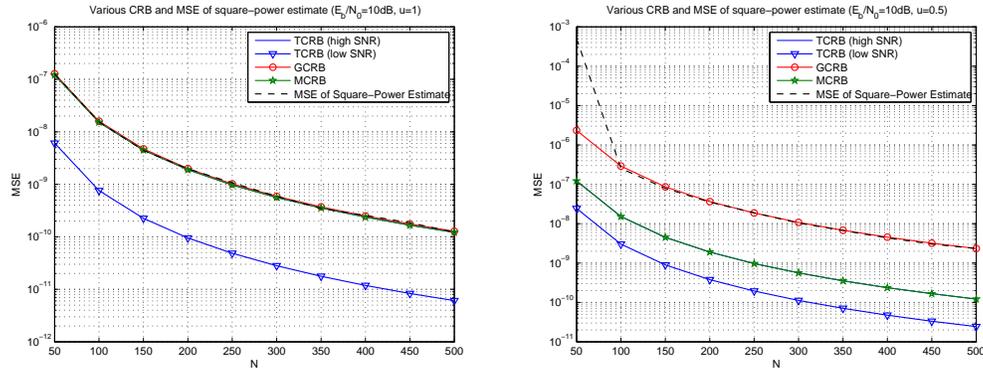


Figure 20: Various CRBs and MSE versus N for $u = 1$ (left) and for $u = 0.5$ (right).

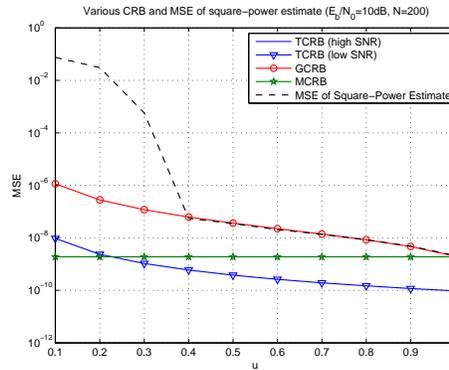


Figure 21: Various CRBs and MSE versus u ($E_b/N_0 = 10\text{dB}$ and $N = 200$).

- DD with hard/soft decision: the hard decision has already been introduced in this tutorial when hard decision has been done after the channel decoding. Hard decision can be replaced with soft decision. As soft decision is usually needed for iterative decoding, we can implement jointly the decoding of turbo or LDPC and the synchronization which leads to the so-called turbo-synchronization concept. More details will be given below.
- LLR maximization : the performance of the system will be better (and so more reliable) if the sync parameters are well chosen. Therefore one can develop sync estimators based cost function dealing with reliable functions. [90].

Here, we will focus on DD with soft decision. The use of soft decision is really interesting at low SNR when synchronization is very difficult. Usually at low SNR, the channel coding requires the use of Turbo-codes or LDPC, namely, of iterative coding. Therefore the next synchronization way is usually called "turbo-synchronization" developed by Vandendorpe-Luise, and others. A nice tutorial treatment is given in [28]. Here, we briefly

provide the main points. For simplicity of exposition, we focus only on phase estimation (assuming the other synchronization parameters are known)

Let $\hat{\phi}^{(n)}$ be the phase estimator at the n -th iteration. The EM algorithm has two basic steps:

$$\begin{aligned} \text{E-Step} & : Q(\phi, \hat{\phi}^{(n)}) = \int p(\tilde{y}|y, \hat{\phi}^{(n)}) \log(p(\tilde{y}|\phi)) d\tilde{y} \\ \text{M-Step} & : \hat{\phi}^{(n+1)} = \arg \max_{\phi} Q(\phi, \hat{\phi}^{(n)}) \end{aligned}$$

where

- \tilde{y} is the *complete* set of data
- y is the *incomplete* set of data.

By considering the complete set of data as the received signal and the transmitted symbols, i.e., $\tilde{y} = [y, \mathbf{s}_N]$, one can obtain that

$$Q(\phi, \hat{\phi}^{(n)}) = \Re \left\{ \sum_{n=0}^{N-1} \check{s}_{n, \hat{\phi}^{(n)}} z(n) e^{-2i\pi\phi} \right\} \quad (23)$$

where

$$\check{s}_{n, \hat{\phi}^{(n)}} = \sum_{p=0}^{P-1} s^{(p)} p(s_n = s^{(p)} | y, \hat{\phi}^{(n)}). \quad (24)$$

Thus the M-step leads to the following solution

$$\hat{\phi}^{(n+1)} = \angle \left(\sum_{n=0}^{N-1} \check{s}_{n, \hat{\phi}^{(n)}} z(n) \right)$$

We thus remark that the phase estimator is very similar to DA or DD estimator but now the symbol is neither known, nor decided, but is replaced with the mean of the a posteriori distribution.

The previous EM approach can be considered with or without coding. The performance will just be different because the a posteriori mean will be more or less accurate compared to the true transmitted symbol (see [48] for non-coded case). In order to connect some previous results, let us consider the non-coded BPSK case. We then have that

$$\check{s}_{n, \hat{\phi}^{(n)}} = p(s_n = 1 | y, \hat{\phi}^{(n)}) - p(s_n = -1 | y, \hat{\phi}^{(n)}) \quad (25)$$

$$= \tanh \left(\frac{1}{2} \log \left(\frac{p(s_n = 1 | y, \hat{\phi}^{(n)})}{p(s_n = -1 | y, \hat{\phi}^{(n)})} \right) \right) \quad (26)$$

$$= \tanh \left(\frac{\Re \left\{ z(n) e^{-2i\pi\hat{\phi}^{(n)}} \right\}}{N_0} \right) \quad (27)$$

Eq. (26) corresponds to the LLR of the symbol, and Eq. (27) corresponds to the standard soft decision on the BPSK symbol. We remark that the EM approach (which is strongly

connected to ML) leads "fortunately" to the same equation of the true ML developed for the non-coded BPSK in Eq. (16) even though the iteration does not mean the same thing. In the EM approach, we iterate inside a data block, whereas in Eq. (16) we iterate at each ongoing sample. The non-coded BPSK based example shows that there exists a strong link between EM and LLR. Nevertheless, this link cannot be extended directly to coded system (whatever the constellation). For more details, we refer the reader to [28].

In terms of complexity, we need to compute Eq. (24). For (coded) BPSK, the BCJR algorithm can be used. For other constellation, usually approximations are done [91, 92]. Notice that we need the soft decision directed approach only at low SNR when the constellation size is in practice small enough which reduces the computational load for obtaining Eq. (24).

Conclusion

In Table 1, we summarize the links between the various estimators that we have discussed and the assumptions under which they work. We also indicate whether each of the problems (1,1',2,2') has a solution.

Problem	Algorithms
1 (DA phase and CFO estimation with null timing)	(9)
1' (DA Timing estimation with null phase and CFO)	(3)
2 (NDA Phase and CFO estimation insensitive to timing)	(20,21,22)
2' (NDA Timing estimation insensitive to phase and CFO)	(19) if $\mathbb{M} = \{0\}$

Table 1: Some algorithms associated with each problem statement.

3. Synchronization for non-flat fading channels

When the channel is frequency-selective, the orthogonal frequency-division multiplexing (OFDM) modulation scheme is the standard of choice. OFDM has been widely employed in various commercial applications that include wireless local area networks (IEEE 802.11a/g/n and HIPERLAN/2), wireless metropolitan area networks (WMAN/WiMax, IEEE 802.16), terrestrial digital audio broadcasting (DAB) and terrestrial digital video broadcasting (DVB) systems in Europe, Multimedia Mobile Access Communications (MMAC) in Japan. The popularity of OFDM stems from its ability to transform a wideband frequency-selective channel to a set of parallel flat-fading narrowband channels, which substantially simplifies the channel equalisation problem. Because of the time-frequency granularity that it offers, OFDM appears to be a natural solution when the available spectrum is not contiguous, for overlay systems, and to cope with issues such as narrowband jamming. In the multiuser context, this granularity also accommodates variable quality-of-service (QoS) requirements and bursty data. A noticeable example of this multiuser application is the combination of OFDM with frequency-division multiple access (FDMA) protocol, i.e., orthogonal frequency-division multiple access (OFDMA), which has become part of the

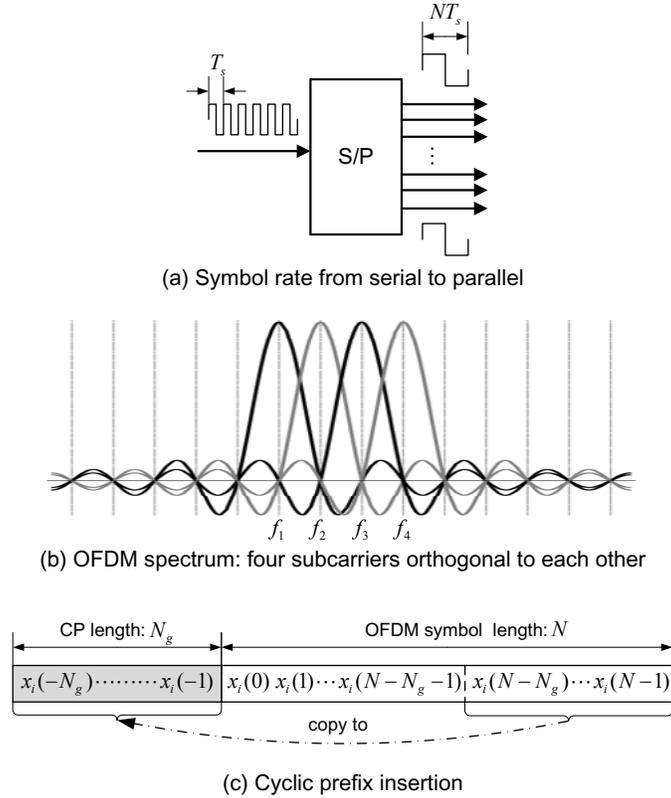


Figure 22: The operational principle of an OFDM system

IEEE 802.16 standards for WMAN. Since subcarriers are allocated to distinct users in a non-overlapping manner, one attractive feature of OFDMA is its capability to mitigate the effects of multiple-access interference (MAI). Another appealing feature of OFDMA is dynamic subcarrier assignment which enables it to optimally allocate system resources such as transmission power and spectrum.

Despite the above-mentioned appealing features, the synchronization task turns out to be a critical issue for OFDM based systems. The synchronization problems of OFDM based systems include timing and frequency synchronization. Timing and frequency offsets come from two sources. One source is the local oscillator frequency mismatch between the transmitter and the receiver, and the other is the Doppler spread due to the relative motion between the transmitter and the receiver. Both timing and frequency synchronization errors introduce extra interference to OFDM systems and result in performance degradation. In addition, timing synchronization may affect the performance of channel estimation [93]. Therefore, effective synchronization is a key to improve the performance of an OFDM based system.

3.1. Signal model and preliminaries

The operational principle of an OFDM system is that the available bandwidth is divided into a large number of subchannels, over each of which the wireless channel can be consid-

ered non-dispersive or flat fading. The original data stream at rate R is split into N parallel data streams, each at rate R/N . The symbol duration, T_s , for these parallel data streams is therefore increased by a factor of N , i.e., $T = NT_s$ as shown in Figure 22.a. Conceptually, each of the data streams modulates a carrier with a different frequency and the resulting signals are transmitted simultaneously. The carriers for each subchannels are made orthogonal to one another, allowing them to be spaced very close together with no overhead. This is shown in Figure 22.b for four carriers. Correspondingly, the receiver consists of N parallel receiver paths. Due to the increased symbol duration, the intersymbol interference (ISI) over each channel is reduced to $\lceil \tau_{\max}/(NT_s) \rceil$ symbols. Thus, an advantage of OFDM is that, for frequency-selective fading channels, the OFDM symbols are less affected by channel fades than are single-carrier transmitted symbols. This is due to the increased symbol duration in an OFDM system. While many symbols during a channel fade might be lost in a single-carrier system, the symbols of an OFDM system can still be correctly detected since only a fraction of each symbol might be affected by the fade. On the other hand, if the channel is time selective, i.e., the channel impulse response varies significantly within the OFDM symbol period, then the channel matrix is no longer Toeplitz and conventional OFDM would fail.

Since multicarrier modulation is based on a block transmission scheme, measures have to be taken to avoid or compensate for interblock interference (IBI), which contributes to the overall ISI. OFDM systems can be categorised by the way they handle IBI. In the most popular systems, a guard time is inserted between consecutive OFDM symbols in the form of a cyclic prefix (CP); i.e., the tail of the OFDM symbol is prefixed as shown in Figure 22.c. The length of the CP is chosen to be larger than the expected delay spread; after proper time synchronization, the receiver discards the CP and thus the IBI is eliminated. Time guarding by zero padding the OFDM symbols has also been proposed in [94, 95]. The issue here is one of turning the transmitter on and off and increased receiver complexity vs. the increased signal-to-noise ratio (SNR) and decreased symbol error rate (SER). Comparisons between cyclic-prefixing and zero-padding OFDM systems may be found in [96]. In this chapter, we focus on CP based OFDM systems only.

The choice of the OFDM parameters is a trade-off between various, often conflicting requirements. The length of the CP is dictated by the delay spread of the channel. Introduction of the CP entails a reduction in rate (or wasted bandwidth), as well as an SNR loss; to minimise these inefficiencies, the number of subcarriers, N , should be large. However, a large number of subcarriers induces high implementation complexity, increased sensitivity to frequency offset and phase noise (since the subcarriers get closer to each other as N increases), and increased peak-to-average power ratios (PAPRs). N is dictated by concerns regarding practical FFT sizes as well as the coherence time of the channel. We will not address the issue of practical choice of OFDM parameters here; we refer the reader to [97] and references therein. In this chapter, we address the crucial issue of timing and frequency offsets estimation.

Subcarrier allocation strategies

Conventionally, all subcarriers are allocated to one specific user in single user OFDM systems. For multiuser OFDM or OFDMA systems, subcarrier allocation strategies are needed.

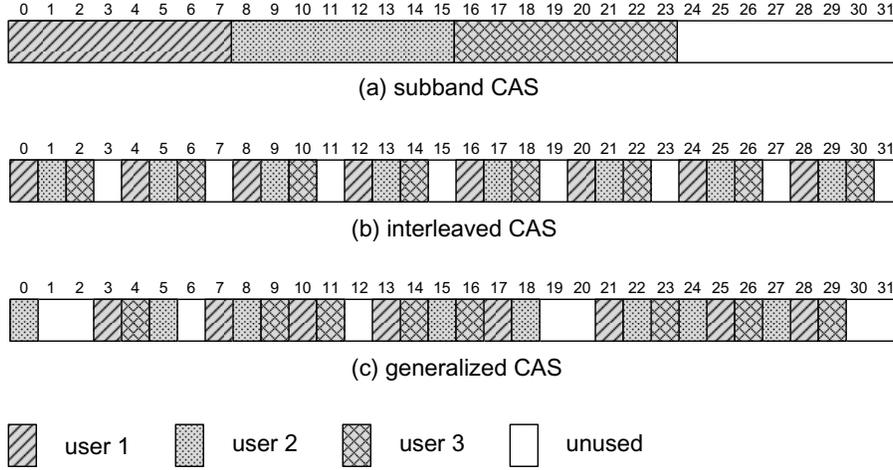


Figure 23: Illustration of subcarrier allocation schemes

Let N and K_u denote the total number of subcarriers and maximum number of active users, respectively. The current number of users M is limited to K_u , i.e., $M \leq K_u$. Let I_i and \mathcal{J}_i indicate the number and indices of subcarriers allocated to the i -th user, respectively; we have that

$$\sum_{i=1}^M I_i \leq N \quad (28)$$

$$\mathcal{J}_i \cap \mathcal{J}_j = \emptyset, \quad i \neq j \quad (29)$$

As shown in Figure 23, generally, there are three subcarrier allocation schemes. For illustration purposes, we set $N = 32$, $M = 3$, $K_u = 4$ and $I_i = 8$ in Figure 23. The *subband carrier allocation scheme* (CAS) is shown in Figure 23(a). A group of I_i adjacent subcarriers is assigned to the i -th user in the subband CAS so that the signal of each user can be separated easily at the base station (BS) through a filter bank. However, subband CAS prevents the possibility of optimally exploiting the channel diversity. A deep fade might hit a substantial number of subcarriers of a given user if they are close together [98]. To reserve some multipath diversity, *interleaved CAS* shown in Figure 23(b) can be adopted for an uplink OFDMA system. The assigned subcarriers of the i -th user are equi-spaced with a distance K_u in interleaved CAS. More dynamic resource allocation and flexibility can be achieved by employing *generalised CAS*, where no strict association between subcarriers and users is required, as illustrated in Figure 23(c). However, generalised CAS will increase the synchronization complexity significantly as shown later.

OFDM transmission

The discrete-time block diagram of a standard downlink OFDMA transmission system is depicted in Figure 24. According to the CAS employed, I_i data symbols for each user and $N - \sum_{i=1}^M I_i$ zeros are assembled into one OFDM symbol as

$$X(k) = \begin{cases} X_i(k), & \text{if } k \in \mathcal{J}_i \\ 0, & \text{otherwise} \end{cases} \quad (30)$$

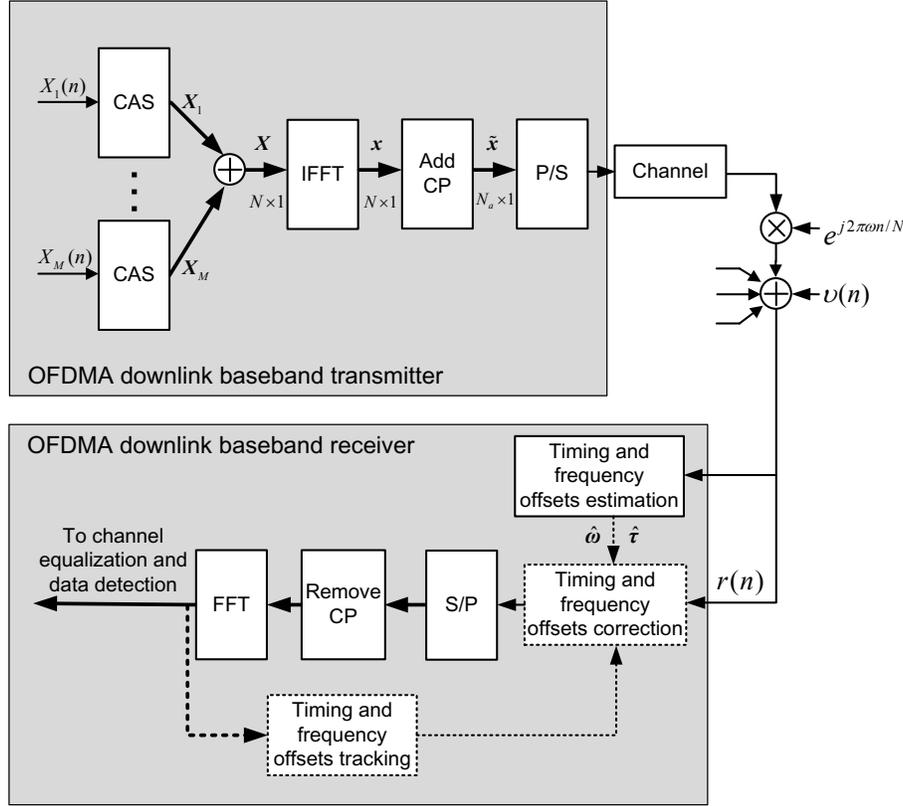


Figure 24: Discrete-time complex baseband representation of downlink OFDMA transmission

After the inverse DFT modulation, each OFDM block is preceded by a CP whose duration is usually longer than the delay spread of the propagation channel, so that IBI can be eliminated at the receiver without affecting the orthogonality of the subcarriers. The time-domain samples after CP insertion can be expressed as

$$\tilde{x}(n) = \begin{cases} x(N+n), & -N_g \leq n \leq -1 \\ x(n), & 0 \leq n \leq N-1 \end{cases} \quad (31)$$

where $x(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k) e^{j2\pi nk/N}$. The $(N_a = N + N_g)$ samples of each block are then pulse shaped, upconverted to the carrier frequency, and transmitted sequentially through the channel.

In this chapter, we model the frequency-selective channel as a finite impulse response (FIR) filter with channel impulse response (CIR) $h = [h(0), \dots, h(L-1)]^T$, where L is the channel order and is determined by the maximum channel delay spread and data sampling rate. In practice, the system is usually designed such that $L \leq N_g \leq N$. We assume that the CIR is time invariant over $N_T \geq 1$ consecutive symbol blocks, but could vary from one set of N_T blocks to the next.

At the receiver, the signal is downconverted to baseband and sampled at the rate of N_a samples per extended OFDM symbol. We will index these received samples by $[-N_g, \dots, N-$

1]. Discarding the samples with indices $n = -N_g, \dots, -1$ is known as discarding the cyclic prefix. In a perfectly synchronised system, the received signal can be written as

$$r(n) = \sum_{l=0}^{L-1} h(l)\tilde{x}(n-l) + v(n) \quad (32)$$

for $n = 0, \dots, N-1$; here $v(n)$ is complex-valued additive white Gaussian noise (AWGN) with variance σ_v^2 . Recall that with the insertion of CP, collected samples $\{r(n)\}_{n=0}^{N-1}$ can be expressed as

$$r = \mathcal{H}x + v \quad (33)$$

where \mathcal{H} is an $(N \times N)$ circulant matrix whose first column is $[h(0), h(1), \dots, h(L-1), 0, \dots, 0]^T$. The circulant matrix \mathcal{H} can be written as $\mathcal{H} = \mathbf{F}^H \mathbf{H} \mathbf{F}$, where \mathbf{F} is the DFT matrix with $[\mathbf{F}]_{m,n} = \frac{1}{\sqrt{N}} e^{-j2\pi mn/N}$ and $\mathbf{H} = \text{diag}\{H(0), H(1), \dots, H(N-1)\}$ with

$$H(k) = \sum_{l=0}^{L-1} h(l) e^{-j2\pi kl/N} . \quad (34)$$

Hence, after performing DFT, the output R can be expressed as

$$R = \mathbf{H}X + Y \quad (35)$$

where $X = [X(0), \dots, X(N-1)]^T$ and $Y = [Y(0), \dots, Y(N-1)]^T$ is again AWGN with covariance matrix $\sigma_v^2 \mathbf{I}$. Since \mathbf{H} is a diagonal matrix, equation (35) indicates that the effect of the frequency-selective channel on the OFDM signal is completely captured by scalar multiplications of the data symbols by the frequency responses of the channel at the subcarrier frequencies. Further, demodulation at the receiver does not colour the additive noise. If none of the channel zeros coincide with an activated subcarrier, maximum likelihood detection of the symbols is straightforward. Zero-forcing and minimum mean square error (MMSE) equalisers can also be applied on a per-carrier basis.

As we mentioned previously, the signal model (35) is only valid for an ideal timing and frequency synchronised system. However, in practical systems, Doppler shifts and instable oscillators result in a carrier frequency offset (CFO) f_0 between the received carrier and the local sinusoids used for signal demodulation. In addition, unknown transmission timing and propagation delay cause the DFT window to be placed in a wrong position at the receiver. This results in a timing error, denoted by t_d , which must properly be compensated to avoid severe performance degradation. Since fractional (normalised to sampling period) timing offsets can be absorbed into the channel, it is a common practice to model the timing offsets as a multiple of the sampling period. Letting ω (a real number) and τ denote the CFO normalised to the subcarrier spacing and the integer part of the timing offset normalised to the sample period, respectively, i.e., $\omega = Nf_0T_s$ and $\tau = \lfloor t_d/T_s \rfloor$, in the presence of timing and frequency offsets, then equation (32) becomes

$$r(n) = e^{j2\pi\omega n/N} \sum_{l=0}^{L-1} h(l)\tilde{x}(n-l-\tau) + v(n) \quad (36)$$

For single timing and frequency offset, the timing and frequency offsets can be estimated and corrected as shown in Figure 24. However, it is a tough task to estimate and compensate multiple timing and frequency offsets. Next, we will review the effects of timing and frequency offsets on the performance of OFDM systems.

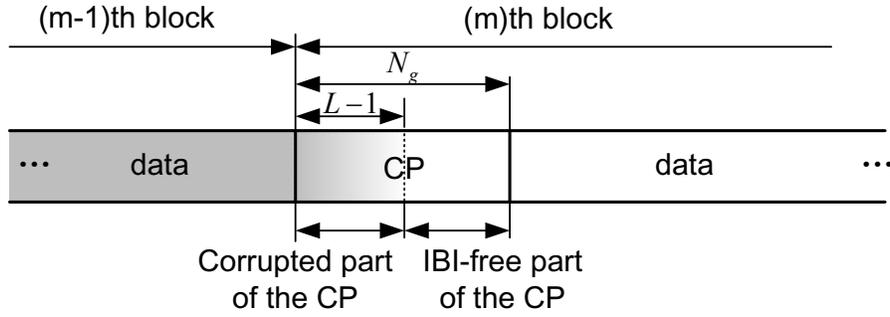


Figure 25: Illustration of the effect of multipath dispersion on CP

Effects of timing errors on system performance

Due to the multipath dispersion, the tail of each received block extends over the first $L - 1$ samples of the successive OFDM block as shown in Figure 25. By inserting a CP which consists of more than $L - 1$ samples, the interference from the previous OFDM block can be readily removed by properly determining the starting point of the OFDM symbol. However, the inaccuracy of timing offset estimation will cause performance degradation. To quantify the effect of timing errors on system performance, we assume perfect frequency synchronization here, i.e., $\omega = 0$.

Since the length of CP is (assumed to be) always larger than the maximum channel delay spread, and using the time-shift property of the Fourier transform, we find that the timing error $\Delta\tau = \hat{\tau} - \tau$ within interval $[-N_g + L - 1, 0]$ only causes a linear phase rotation across the subcarriers as

$$R(k) = e^{j2\pi k \Delta\tau / N} H(k)X(k) + Y(k) \quad (37)$$

The effect of this timing error can be readily compensated by the channel equaliser. On the other hand, if the timing error is outside interval $[-N_g + L - 1, 0]$, samples from adjacent OFDM blocks not only cause IBI, but also result in a loss of orthogonality among subcarriers which generates inter-carrier interference (ICI). A comprehensive mathematical analysis of the effects of timing errors is discussed in [93] and [99]. For the second timing error case, the received signal after DFT can be written as

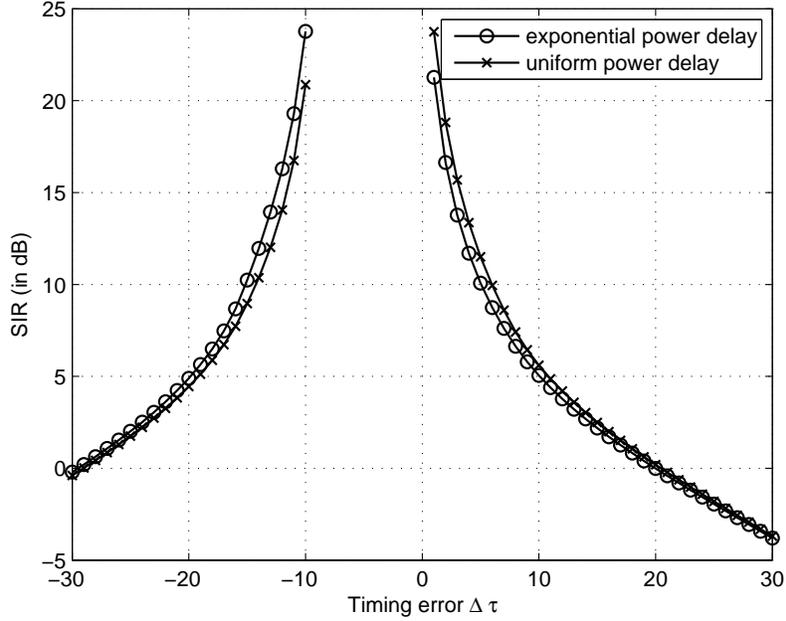
$$R(k) = \frac{\alpha_d}{N}(k) e^{j2\pi k \Delta\tau / N} H(k)X(k) + \gamma(k) + Y(k) \quad (38)$$

where

$$\alpha_d(k) = \begin{cases} (1 - e^{j2\pi k d / N}) / (1 - e^{j2\pi k / N}), & k \neq 0 \\ N - d, & k = 0 \end{cases} \quad (39)$$

$$d = \begin{cases} \Delta\tau, & \text{if } \Delta\tau > 0 \\ \max\{L - 1 - (N_g + \Delta\tau), 0\}, & \text{if } \Delta\tau \leq 0 \end{cases} \quad (40)$$

and $\gamma(k)$ is the combination of IBI and ICI which is defined as [99]. Following the same lines as derived in [99], the signal-to-interference (SIR) in the presence of timing errors can

Figure 26: SIR versus timing error $\Delta\tau$

be expressed as

$$\text{SIR} = \frac{(N-d)^2}{d(2N-d) - 2\frac{N-d}{\sigma_h^2}\beta(d)} \quad (41)$$

where $\sigma_h^2 = \sum_{l=0}^{L-1} \sigma_{h(l)}^2$ and

$$\beta(d) = \begin{cases} \sum_{m=0}^{d-1} \sum_{l=m+1}^{L-1} \sigma_{h(l)}^2, & \Delta\tau > 0 \\ \sum_{m=0}^{d-1} \sum_{l=0}^{N_g+\Delta\tau+m-1} \sigma_{h(l)}^2, & \Delta\tau \leq 0 \end{cases} \quad (42)$$

Figure 26 shows the SIR versus timing error $\Delta\tau$ for $N = 64$. Both the exponential power delay profile, i.e., $E\{|h(l)|^2\}_{l=0}^{L-1} = C \exp(-0.2l)$ where C is a scalar factor that ensures that the total energy of the channel taps is normalised to unity, and the uniform power delay profile, i.e., $E\{|h(l)|^2\}_{l=0}^{L-1} = 1/L$ profile, have been tested. The length of the CP and CIR are set to 16 and 8, respectively, i.e., $N_g = 16$ and $L = 8$. As discussed previously, the timing error within interval $[-N_g + L - 1, 0]$, which is $[-9, 0]$ for our simulation setup, will not cause ICI and IBI. The timing error outside the interval $[-9, 0]$ results in a significant SIR loss, especially for positive timing errors. To keep the SIR degradation within a tolerable level, accurate timing offset estimation is necessary.

Effects of carrier frequency offsets on system performance

As we mentioned in a previous section, CFO is caused by Doppler shifts and mismatched oscillators. In general, the CFO can be several times the subchannel spacing. Thus, the CFO is usually divided into an integer part and a fractional part by normalising to the subcarrier spacing. Assuming perfect timing synchronization and using the frequency-shift property of the Fourier transform, the received signal under an integer valued CFO ω can be expressed as

$$R(k) = H((k - \omega) \bmod N)X((k - \omega) \bmod N) + Y(k) \quad (43)$$

It can be seen from the above equation that the integer valued CFO causes a circular shift of the transmitted symbols, but does not cause ICI; i.e., the orthogonality of the subcarriers is maintained. The fractional part, however, causes ICI. Assuming that CFO ω is a fractional value, the received signal can be written as

$$R(k) = \sum_{n=0}^{N-1} H(n)X(n)f(\omega + n - k) + Y(k) \quad (44)$$

where

$$f(n) = \frac{\sin(\pi n)}{N \sin(\pi n/N)} e^{j\pi n(N-1)/N} \quad (45)$$

Equation (44) can be re-written as

$$R(k) = H(k)X(k)f(\omega) + \gamma(k) + Y(k) \quad (46)$$

where $\gamma(k) = \sum_{n=0 \neq k}^{N-1} H(n)X(n)f(\omega + n - k)$ is the zero-mean ICI term with power $\sigma_\gamma^2(\omega)$.

After some manipulations as shown in [100], we have

$$\sigma_\gamma^2(\omega) = \sigma_x^2(1 - |f(\omega)|^2)$$

Thus, the SIR can be expressed as

$$\text{SIR} = \frac{|f(\omega)|^2}{1 - |f(\omega)|^2} \quad (47)$$

The SIR versus CFO ω for $N = 64$ is shown in Figure 27. Notice that the SIR decreases rapidly as the CFO increased. Again, to keep the SIR degradation to a tolerable level, effective CFO estimation and compensation methods are required. More precise techniques for computing the SNR loss due to CFO can be found in [101].

3.2. Downlink OFDMA

As shown in Figure 28, the BS broadcasts training sequences followed by data blocks to the potential users and each user operates independently in downlink OFDMA transmission. Thus, the synchronization problem in downlink OFDMA is similar to that of single-user OFDM systems. Generally, synchronization can be divided into a coarse acquisition phase and a fine tracking phase. In this section, we provide a brief survey of existing synchronization techniques in the downlink OFDMA scenario.

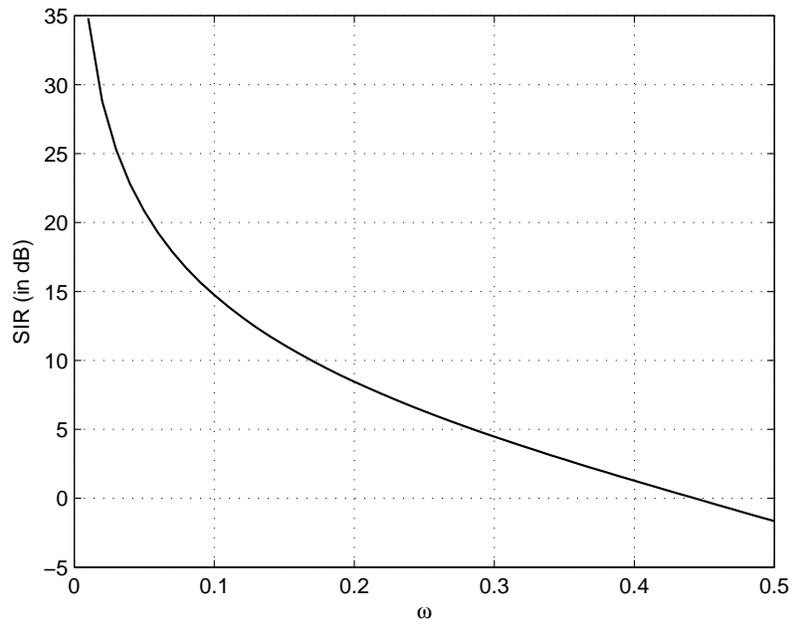
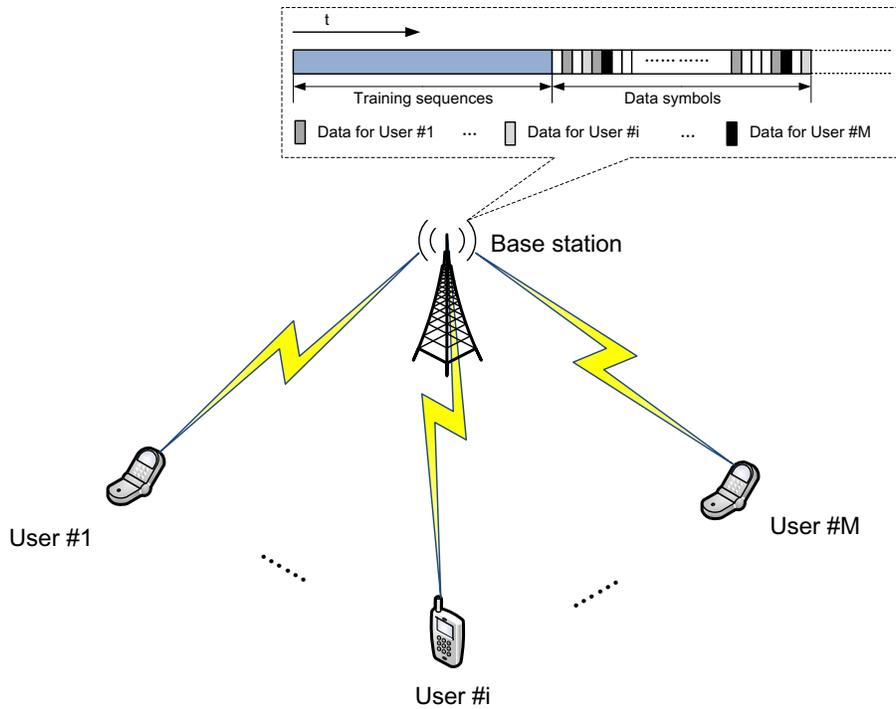
Figure 27: SIR versus fractional carrier frequency offset ω 

Figure 28: Downlink OFDMA representation

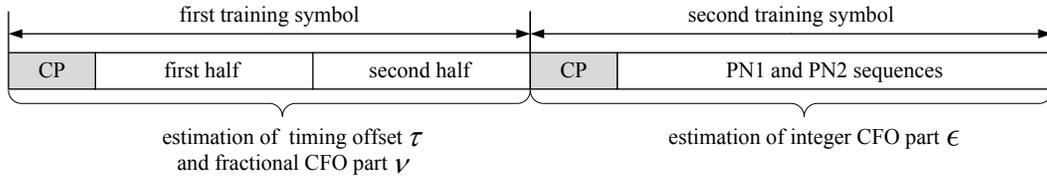


Figure 29: Training symbols of S&C algorithm

Coarse synchronization methods

The coarse synchronization task typically has two sub-tasks, i.e., finding the start of an OFDM frame over an approximate range of sample values and aligning the local oscillator of the receiver to the received carrier frequency. Coarse timing acquisition is usually the first task performed in the synchronization procedure. In practice, the CFO is assumed to be completely unknown at this stage. Hence, the orthogonality of the subcarriers may not be retained to provide a useful post-FFT signal. Consequently, coarse timing acquisition is obtained in the time domain. In coarse frequency synchronization, the usual approach is to decompose the CFO into a fractional part plus an integer part. Pre-FFT or post-FFT methods may be adopted to estimate the CFO.

Depending on the system requirements, coarse timing and frequency acquisition can be carried out by exploiting either the repeated cyclic prefix [102, 103, 104] or specially designed training sequences (preambles) [105, 106, 107, 108]. Exploiting the correlation of the CP, the CP based algorithms can work blindly without the overhead of an explicit training sequence. However, as standardised in many commercial systems, reliable coarse acquisition methods for frequency-selective channels are based on a training sequence with a repetitive structure. The motivation behind this idea is that the repetitive property is preserved after propagation through a multipath channel, except for a phase rotation due to the CFO. In this chapter, we consider training-sequence based synchronization in the following.

Second-order statistics based methods

There are basically two methods for training-sequence based coarse synchronization, i.e., first-order statistics based and second-order statistics based methods. The latter class was first proposed by Moose in [109] and further studied by Schmidl and Cox (S&C) in [105], where two identical slots with length of $N/2$ were used in the first training symbol as shown in Figure 29. The CFO normalised to the subcarrier spacing is decomposed into two parts as

$$\omega = \nu + 2\epsilon \quad (48)$$

where $\nu \in (-1, 1]$ and ϵ is an integer. To generate the repetitive-structure of the first training symbol, we can simply transmit a pseudonoise (PN) sequence on the even subcarriers, while zeros are used on the odd subcarriers. For the second training symbol, a PN sequence PN1 is transmitted on the odd subcarriers which may be employed in channel estimation; and a differentially-modulated PN sequence PN2 deployed on the even subcarriers is used for the integer CFO estimation. Let $\{r(n)\}_{n=0}^{N-1}$ denote the received signals, we have

$$r(n) = z(n) + v(n) \quad (49)$$

where

$$z(n) = e^{j2\pi\omega n/N} \sum_{l=0}^{L-1} h(l)x(n-\tau-l) \quad (50)$$

$$x(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N/2-1} X(2k)e^{-j4\pi kn/N} \quad (51)$$

where τ , an integer, is the normalised nonfractional part of the timing offset, ω is the CFO normalised to the subcarrier spacing, $x(n)$ and $X(k)$ are respectively the transmitted time domain and frequency domain training sequences, $h(l)$ is the l th tap of the CIR and $v(n)$ is an AWGN with variance σ_v^2 . It can be easily verified that the first and second halves of the received signal can be expressed as

$$r(n) = z(n) + v(n), \quad \tau \leq n \leq \tau + \frac{N}{2} - 1 \quad (52)$$

$$r(n + \frac{N}{2}) = e^{j\pi v} z(n) + v(n + \frac{N}{2}), \quad \tau \leq n \leq \tau + \frac{N}{2} - 1 \quad (53)$$

Exploiting the correlation between first and second halves, the S&C timing estimator can be expressed as

$$\hat{\tau} = \arg \max_{\tilde{\tau}} \Lambda(\tilde{\tau}) \quad (54)$$

where

$$\Lambda(\tilde{\tau}) = \frac{\left| \sum_{n=\tilde{\tau}}^{\tilde{\tau}+N/2-1} r^*(n)r(n+\frac{N}{2}) \right|^2}{\left(\sum_{n=\tilde{\tau}}^{\tilde{\tau}+N/2-1} |r(n+\frac{N}{2})|^2 \right)^2}$$

Moreover, assuming perfect timing synchronization, the estimate of fractional part CFO v can be obtained as

$$\hat{v} = \frac{1}{\pi} \arg \left\{ \sum_{n=\tau}^{\tau+N/2-1} r^*(n)r(n+\frac{N}{2}) \right\} \quad (55)$$

In practice, the timing offset τ in eq. (55) can be replaced by its estimated value $\hat{\tau}$ given in eq. (54). If the normalised CFO can be guaranteed to be less than 1, the second training symbol would not be needed. Otherwise, the second training symbol and a post-FFT method can be adopted to estimate the integer part of the CFO, ε , as we describe next.

After compensating the fractional offset by multiplying the two training symbols by $e^{-j2\pi\hat{v}n/N}$, the FFT output of two training symbols, denoted as $R_1(k)$ and $R_2(k)$, can be expressed as

$$R_1(k) = Z_1(k) + W_1(k) \quad (56)$$

$$R_2(k) = Z_2(k) + W_2(k) \quad (57)$$

for $k = 0, \dots, N-1$ and

$$Z_1(k) = H((k-2\varepsilon)_{\text{mod}N})X_1((k-2\varepsilon)_{\text{mod}N}) \quad (58)$$

$$Z_2(k) = e^{j\phi} H((k-2\varepsilon)_{\text{mod}N})X_2((k-2\varepsilon)_{\text{mod}N}) \quad (59)$$

where $\phi = 4\pi\varepsilon(N + N_g)/N$. Let $d(k) = \frac{X_2(k)}{X_1(k)}$ represent the differentially-modulated PN sequence on the even subcarriers of the second training symbol, we have

$$Z_2(k) = e^{j\phi} d((k - 2\varepsilon)_{\text{mod}N}) Z_1(k) \quad (60)$$

for even k . Thus, the estimator of the integer part of the CFO, ε , can be expressed as

$$\hat{\varepsilon} = \arg \max_{\tilde{\varepsilon}} \Psi(\tilde{\varepsilon}) \quad (61)$$

where

$$\Psi(\tilde{\varepsilon}) = \frac{\left| \sum_{k \text{ even}} R_1^*(k) d^*((k - 2\tilde{\varepsilon})_{\text{mod}N}) R_2(k) \right|^2}{2 \left(\sum_{k \text{ even}} |R_2(k)|^2 \right)^2} \quad (62)$$

and integer $\tilde{\varepsilon}$ varies over the range of possible frequency offsets. Then, the frequency offset estimate would be $\hat{\omega} = \hat{\nu} + 2\hat{\varepsilon}$.

The S&C timing estimator (54) is expected to capture a peak when the correlation window is perfectly aligned with the received training sequence. Unfortunately, as shown in Figure 30, the timing metric of the S&C estimator exhibits a "plateau" which reduces the acquisition accuracy significantly. To obtain a steeper timing metric trajectory, many training patterns were proposed in [106, 107]. In [107], a training of the form $[\mathbf{B}, \mathbf{B}, -\mathbf{B}, \mathbf{B}]$ in time domain was proposed by Shi and Serpedin (S&S). Arranging the received N samples $[r(\tilde{\tau}) \cdots r(\tilde{\tau} + N - 1)]^T$ into four parts as $\{r_i(\tilde{\tau}) = [r(iN/4 + \tilde{\tau}) \cdots r((i+1)N/4 + \tilde{\tau} - 1)]^T\}_{i=0}^3$, the S&S timing estimator can be expressed as

$$\hat{\tau} = \arg \max_{\tilde{\tau}} \Lambda(\tilde{\tau}) \quad (63)$$

where

$$\Lambda(\tilde{\tau}) = \frac{\sum_{i=0}^2 |P_i(\tilde{\tau})|}{\frac{3}{2} \sum_{i=0}^3 \|r_i(\tilde{\tau})\|^2}$$

and

$$\begin{aligned} P_0(\tilde{\tau}) &= r_0^H(\tilde{\tau})r_1(\tilde{\tau}) - r_1^H(\tilde{\tau})r_2(\tilde{\tau}) - r_2^H(\tilde{\tau})r_3(\tilde{\tau}) \\ P_1(\tilde{\tau}) &= r_1^H(\tilde{\tau})r_3(\tilde{\tau}) - r_0^H(\tilde{\tau})r_2(\tilde{\tau}) \\ P_2(\tilde{\tau}) &= r_0^H(\tilde{\tau})r_3(\tilde{\tau}) \end{aligned}$$

Since the training symbol is divided into four parts, we find that the CFO causes a phase shift of $\pi\omega/2$ in each part for a flat fading channel. Thus, the CFO estimator of S&S algorithm for practical systems can be expressed as

$$\hat{\omega} = \frac{2}{\pi} \arg \{P_0(\hat{\tau})\} \quad (64)$$

Compared to the S&C CFO estimator (55), we can see that the acquisition range of (64) increased to $[-2, 2)$.

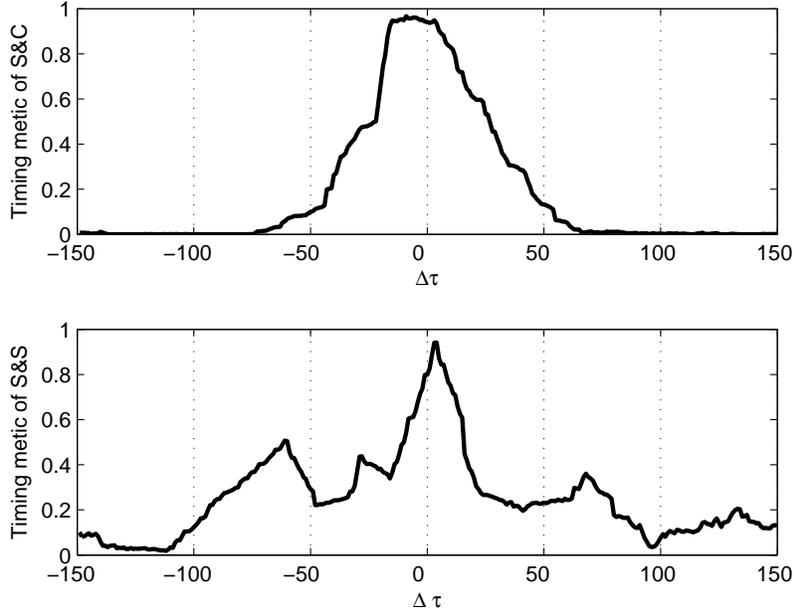


Figure 30: Timing metrics for S&C and S&S estimators, $N = 128$, $N_g = 16$, $L = 16$, SNR=20dB

Both timing metrics of the S&C and S&S timing estimators, $\Lambda(\tilde{\tau})$, are illustrated in Figure 30, where $\Delta\tau = \tilde{\tau} - \tau$. The results are obtained under the exponential power delay profile channel introduced previously and the SNR is defined as σ_x^2/σ_v^2 . We can see that the "plateau" present in the S&C estimator is significantly reduced in the S&S estimator. As pointed out in [106], a steeper timing metric trajectory can be obtained by increasing the number of repetitive slots.

Although the correlation method adopted in S&C and S&S estimators has low computational complexity, those estimators will exhibit a floor effect since they are still based on second-order statistics of the received signal. As shown in [110], much more accurate timing and frequency estimation can be achieved by using the first-order statistics, at the expense of a slight increase in implementation complexity.

First-order statistics based methods

Using the signal model (49), the mean of the received signal is given by

$$E\{r(n)\} = e^{j2\pi\omega n/N} \sum_{l=0}^{L-1} h(l)x(n-\tau-l), n = \tau - N_g, \dots, \tau + N + L - 1 \quad (65)$$

Let $r_{\tilde{\tau}} = [r(\tilde{\tau}), \dots, r(\tilde{\tau} + N - 1)]^T$ and $E\{r_{\tilde{\tau}}\} \triangleq \mu = \Gamma(\omega)\mathbf{X}h$, where $\Gamma(\omega) = \text{diag}\{1, \dots, e^{j2\pi\omega(N-1)/N}\}$. \mathbf{X} is a circulant matrix whose first column is x .

The variance of the received signal is

$$\text{var}\{r(n)\} = \begin{cases} \sigma_v^2, & n = 0, \dots, \tau + N - 1 \\ > \sigma_v^2, & n > \tau + N - 1 \end{cases} \quad (66)$$

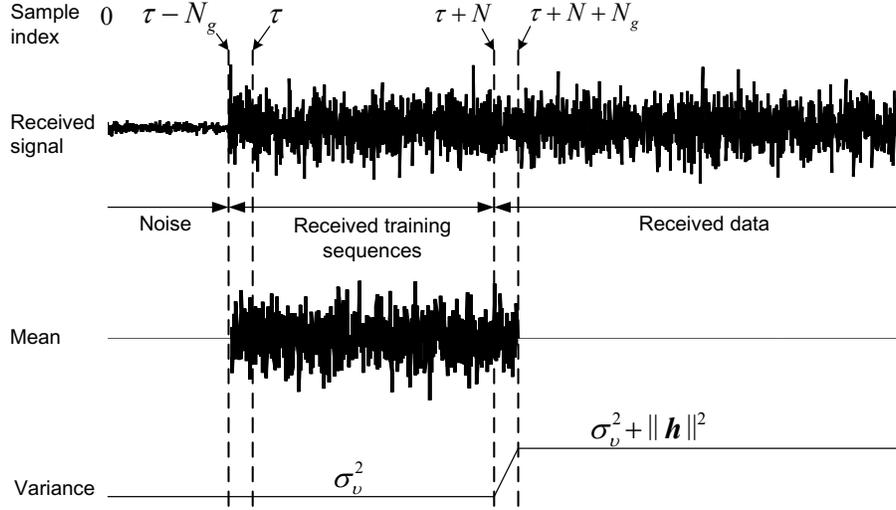


Figure 31: Real part of one realisation of the received signal as well as the corresponding instantaneous mean and variance

Thus, the variance of the received signal is minimum during the noise-only period, which precedes reception of the frame, and during the reception of the training block. During data reception, it is equal to $\sigma_v^2 + \|h\|^2$, where we assume that data symbols have unit power, without loss of generality. Figure 31 illustrates these observations. If μ were known, τ could be estimated by minimising the Euclidean distance between $r_{\tilde{\tau}}$ and μ over $\tilde{\tau}$. Since this is not the case and in order to avoid the noise-only period, [110] proposed two estimators which are obtained by minimising the following modified versions of the nonlinear least-squares (NLLS) criterion:

$$C_1(\tilde{\tau}, \tilde{\omega}, \tilde{h}) = \frac{\|r_{\tilde{\tau}} - \tilde{\mu}\|^2}{\|r_{\tilde{\tau}}\|^2} \quad (67)$$

and

$$C_2(\tilde{\tau}, \tilde{\omega}, \tilde{h}) = \|r_{\tilde{\tau}} - \tilde{\mu}\|^2 - \|r_{\tilde{\tau}}\|^2 \quad (68)$$

over $\tilde{\tau}$, $\tilde{\omega}$ and \tilde{h} , where $\tilde{\mu}$ is obtained as in μ after replacing h and ω by \tilde{h} and $\tilde{\omega}$, respectively.

Both the normalisation factor in (67) and the second term of the RHS of (68) guarantee the uniqueness of the solution, i.e., avoid the noise-only (i.e., $\tilde{h} = 0$) solution. Indeed, in the noise-only period, the minima of $E\{C_1\}$ and $E\{C_2\}$, which are obtained with $\tilde{h} = 0$, are (approximately) one and zero, respectively. During data reception, the minima of $E\{C_1\}$ and $E\{C_2\}$ are again obtained with $\tilde{h} = 0$ and are also approximately equal to one and zero, respectively. During the reception of the training sequence, the minima of the $E\{C_1\}$ and $E\{C_2\}$ are (approximately) $1/(1 + \text{SNR})$ and $-\|h\|^2$, respectively, where SNR is defined as $\text{SNR} = \|r_{\tau}\|^2 / (N\sigma_v^2)$. These minima are smaller than those obtained in the noise-only and data transmission periods. Hence, when the processed signal contains the received preamble, the minima of $E\{C_1\}$ and $E\{C_2\}$ are achieved if and only if $\tilde{\tau} = \tau$ and $\tilde{\mu} = \mu$. Figure 31 illustrates some of these results.

Since the statistical expectation of C_1 and C_2 are unknown, only estimates of the unknown parameters can be obtained by minimising C_1 and C_2 themselves. The obtained es-

timators were referred to as modified nonlinear least squares estimators (MNLLS) in [110], since they combine the NLLS estimation method and detection. The above optimisation problems can be simplified by noting that C_1 and C_2 are quadratic in \tilde{h} . Thus, closed-form expressions for \hat{h}_1 and \hat{h}_2 can be obtained as

$$\hat{h}_i = \mathbf{X}^\dagger \Gamma^H(\hat{\omega}_i) r_{\hat{\tau}_i}, \quad i = 1, 2. \quad (69)$$

Substituting the above estimates for \tilde{h} in C_1 and C_2 , the equivalent criterion is to maximise

$$C_i''(\tilde{\tau}, \tilde{\omega}) = g_i(\tilde{\tau}) r_{\tilde{\tau}}^H \Gamma(\tilde{\omega}) \Pi_X \Gamma^H(\tilde{\omega}) r_{\tilde{\tau}}, \quad i = 1, 2 \quad (70)$$

where $g_1(\tilde{\tau}) = \|r_{\tilde{\tau}}\|^{-2}$ and $g_2(\tilde{\tau}) = 1$, and $\Pi_X = \mathbf{X}(\mathbf{X}^H \mathbf{X})^{-1} \mathbf{X}^H$, which is a fixed matrix and can thus be precomputed and stored at the receiver.

The above optimisation problems are two-dimensional. Although they are discrete in one dimension, they are still computationally challenging to solve. To reduce the computational complexity, timing acquisition using C_i'' is performed by ignoring the CFO-related terms. It was shown that a coarse but closed-form estimate of the CFO for each timing offset candidate is good enough to (nearly) obtain $\hat{\tau}_1$ and $\hat{\tau}_2$. To obtain the CFO estimate, the repetitive structure of the training block and the second-order statistic-based method in [105] were adopted to estimate the fractional part of ω , i.e., ν . The estimate of the integer part of ω , $\hat{\varepsilon}$ and timing offset $\hat{\tau}_i$ are given by

$$\{\hat{\tau}_i, \hat{\varepsilon}_i\} = \arg \max_{\tilde{\tau}, \tilde{\varepsilon}} g_i(\tilde{\tau}) r_{\tilde{\tau}}^H \Gamma(\tilde{\varepsilon} + \tilde{\nu}) \Pi_X \Gamma^H(\tilde{\varepsilon} + \tilde{\nu}) r_{\tilde{\tau}} \quad (71)$$

where candidate values for

$\tilde{\tau}$ are in $(-Q + 1, Q - 1)$, and $\tilde{\nu}$ is given by [105]

$$\tilde{\nu} = \frac{1}{\pi} \arg \left\{ \sum_{n=\tilde{\tau}}^{\tilde{\tau}+N/2-1} r^*(n) r(n+N/2) \right\} \quad (72)$$

Although the optimisation problem (71) is two-dimensional, it is discrete and the possible values of ε dictated by Q may be small in practice. It is worth pointing out that the optimisation problems can be reduced to one-dimensional problems if the preamble is made of Q repetitive slots, since in this case, a closed-form estimate of ω can be obtained (see [111] and [112]). However, the performance of timing acquisition in this case becomes similar to that of existing methods, unless $N \gg 2Q(L+1)$.

Once τ and ε are estimated, estimates of ω are obtained as $\hat{\omega}_i = \hat{\nu} + \hat{\varepsilon}$, where $\hat{\nu}$ is given by (72) with $\tilde{\tau} = \hat{\tau}$. A more accurate estimate can be obtained by maximising $C_i''(\hat{\tau}, \tilde{\omega})$ over $\tilde{\omega}$ after initialising with $\hat{\omega}_i$. Since this results in the optimisation of a continuous-valued variable, ω , it may not be appropriate in practice, especially for the downlink. Moreover, simulations have shown that the performance improvement is not significant.

Finally, as a product of the above synchronization method, estimates of the channel can be obtained from (69) after replacing τ and ω by the above estimates.

To further reduce the complexity of the computation of the cost functions C_i'' , the projection matrix Π_X can be replaced by $(1/N)\mathbf{X}\mathbf{X}^H$ obtained by approximating $\mathbf{X}^H \mathbf{X}$ by $N\mathbf{I}$, since $N/2 > L + 1$, and using the law of large numbers. Using this, C_2'' is obtained as

the squared L_2 -norm of a vector and C_1'' is its normalised version, and the corresponding estimates of τ_i and ε_i are given by

$$\{\check{\tau}_i, \check{\varepsilon}_i\} = \arg \max_{\check{\tau}, \check{\varepsilon} \in (-Q/2, Q/2)} N^{-1} g_i(\check{\tau}) \|\mathbf{X}^H \Gamma^H (\check{\mathbf{v}} + \check{\varepsilon}) r_{\check{\tau}}\|^2 \quad (73)$$

Note that the computation of $\mathbf{X}^H \mathbf{y}$ requires only changing the sign of some of the elements of the vector \mathbf{y} and additions. Also, note that the computation of the simplified version of C_2'' is simpler than that of C_1'' since in the former, $\|r_{\check{\tau}}\|^2$ needs to be computed. However, the normalisation in C_1'' is desirable for reducing the complexity of the search procedure for real-time implementation.

Fine synchronization methods

In certain applications, due to unstable oscillators or high-mobility environment, the coarse estimates may be inadequate for the entire frame. Thus, timing and frequency tracking are required to compensate the short-term variations produced by oscillator drifts and/or time-varying Doppler shifts.

Fine timing synchronization can be achieved through correlation either in the time domain [113, 114] or in the frequency domain [115, 116, 117, 118]. A timing tracking scheme based on time-domain PN-sequence correlation and a weighted time correlation scheme exploiting the redundancy in both the cyclic prefix and available pilot symbols were shown to provide better performance than repetitive-structured OFDM symbols in [113] and [114], respectively. Frequency-domain based schemes model the timing error as part of the CIR vector. This is motivated by the fact that the estimated CIR shifts cyclicly with respect to the FFT-window. Since the frequency-domain based schemes can resolve channel multipaths effectively, they generally provide better performance than the time-domain schemes in multipath fading channels. There are several approaches to update the coarse timing estimate in frequency-domain based tracking algorithms. One method is to locate the peak position of the estimated CIR \hat{h} [116]. Many modified peak-finding versions were studied in [106], to take into account both the effects of noise, as well as the fact the first peak may not be the largest peak in the CIR. Another method which maximises the energy window of the channel estimation has been investigated in [115, 103]. Exploiting the timing information embedded in pilot-aided channel estimation, timing can be estimated without a specific training sequence, as was shown in [119] and [120].

Similar to fine timing synchronization, frequency tracking can be performed either in the time domain [121, 102] or in the frequency domain [122, 123]. In [121, 102], the correlation between the CP and the last N_c samples of each block is exploited to estimate the residual frequency offset. The residual CFO was tracked using the temporal correlation in the data-aided post-FFT stage and the frequency domain channel estimate was adopted to deduce the weights for a weighted-least-squares CFO estimator in [122] and [123], respectively. Several blind CFO estimation methods, e.g., [124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134], also can be employed during the tracking step without increasing synchronization overhead.

Simulation results for downlink OFDMA systems

We consider an downlink OFDMA system with a total of $N = 128$ subcarriers. In the simulation, we assume that the CIR length is $L = 16$, and a cyclic prefix of length $N_g = 16$. The channel taps $h(l)$ are uncorrelated zero-mean Gaussian random variables with exponential power delay profile $E\{|h(l)|^2\}_{l=0}^{L-1} = C \exp(-0.2l)$, where C is a scalar factor that ensures that the total energy of the channel taps is normalised to unity. Correspondingly, the SNR of received signal is equal to σ_x^2/σ_v^2 , where σ_x^2 is the power allocated to each subcarrier. Since only the fractional part of the CFO can be estimated by the S&S algorithm in [107], we generate ω randomly from the interval $[-0.5, 0.5]$ and generate a new random channel for each Monte Carlo run. Moreover, as the integer part of the CFO is zero, the second training symbol shown in Figure 29 is unnecessary and we set $Q = 1$ for the first-order statistics based methods.

The results of timing and frequency estimates are calculated using 20000 Monte Carlo runs. The figures show that the first-order statistics based methods significantly outperform the second-order statistics based methods in terms of timing and CFO estimation. For timing estimation, MNLLS1 outperforms MNLLS2. The simplified MNLLS algorithms yield the same timing estimation performance. In Figure 33, the probability of exact timing refers to the probability that the associated algorithm identifies τ without error. For comparison, Figure 34, we also show the Cramér-Rao bound (CRB) for the case where timing is perfect. Using the method in [135], the CRB is found to be

$$\text{var}\{\hat{\omega}\} \geq \text{CRB}(\hat{\omega}) = \frac{2\sigma_v^2}{N\pi^2\sigma_x^2} \quad (74)$$

Notice that the performance of first-order statistics based methods is close to CRB. The simplified MNLLS methods have the same CFO estimation performance merits as the MNLLS methods. From the simulation results, we can see that the S&S algorithm achieves more accurate timing estimation performance than the S&C algorithm at the price of a decrease in the CFO estimation performance. As we discussed previously, the gains of first-order statistics based methods come along with the increasing complexity. For real-time implementation, it is important to set a threshold on the synchronization criterion so that complexity can be reduced by pruning the set of timing candidates. Details of implementation issues of first-order statistics based methods can be found in [110].

4. Multiuser synchronization

4.1. Uplink signal model and synchronization policy

We consider an uplink OFDMA system where M active users simultaneously communicate with the BS as shown in Figure 35. The users' data streams are assembled into OFDM symbols according to the CAS employed.

Let $r(n)$ denote the signal received at the BS; we have that

$$r(n) = \sum_{i=1}^M z_{(i)}(n) + v(n) \quad (75)$$

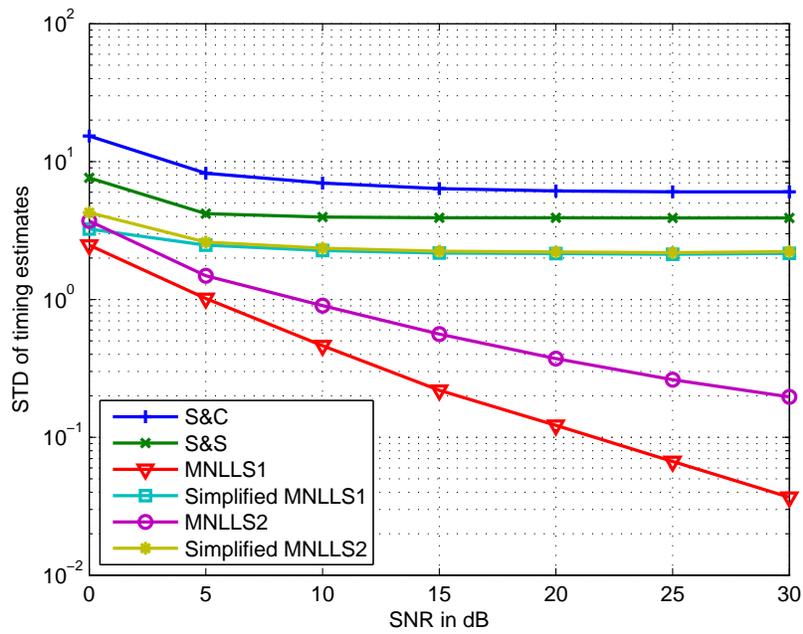


Figure 32: Standard deviation of timing estimates for downlink OFDMA

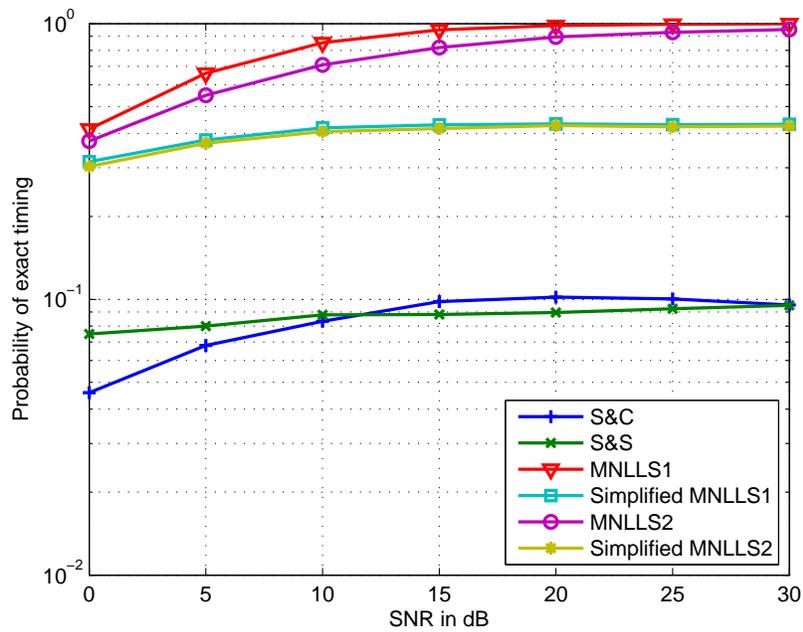


Figure 33: Probability of exact timing acquisition for downlink OFDMA

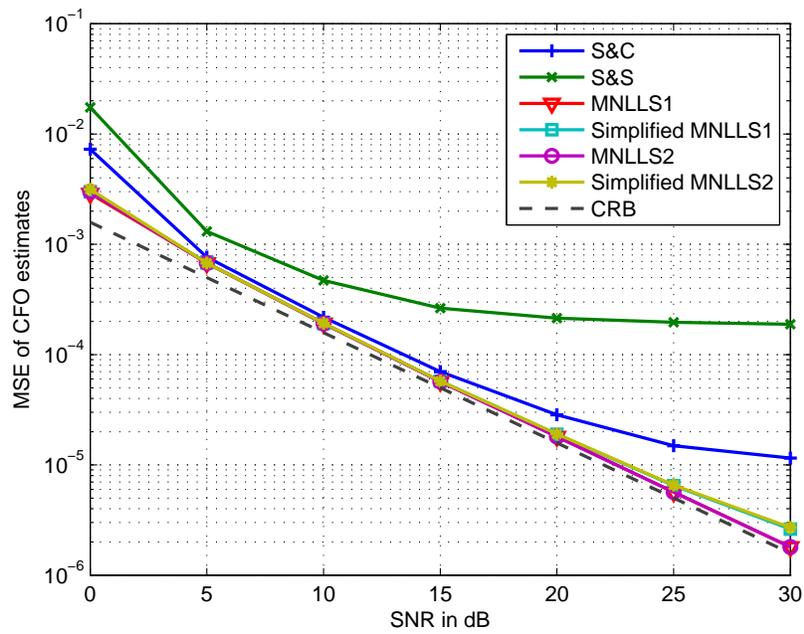


Figure 34: Mean square error of CFO estimates for downlink OFDMA

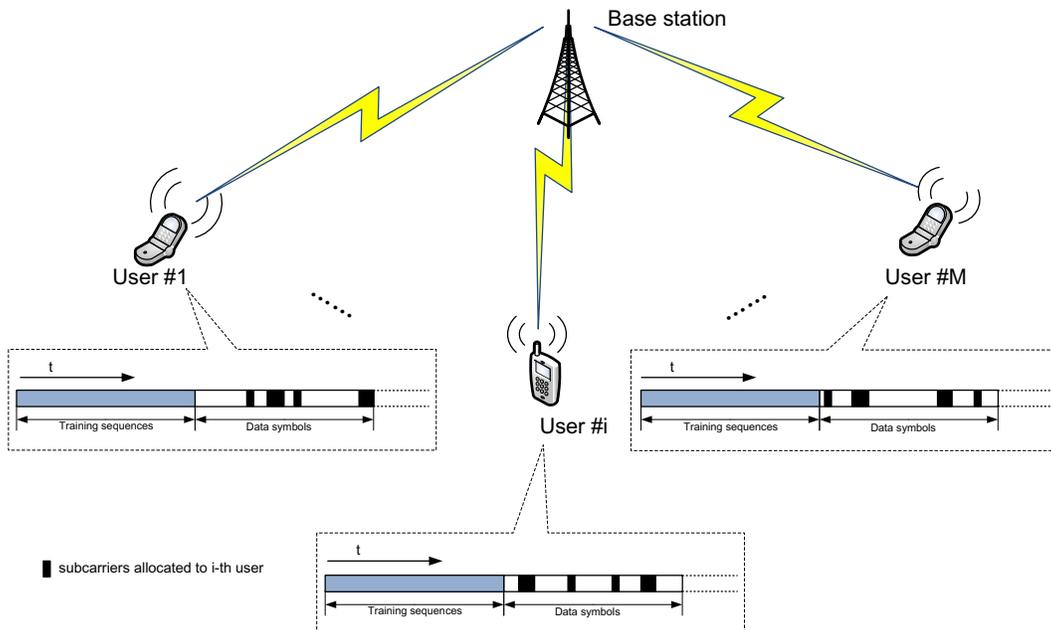


Figure 35: OFDMA uplink representation

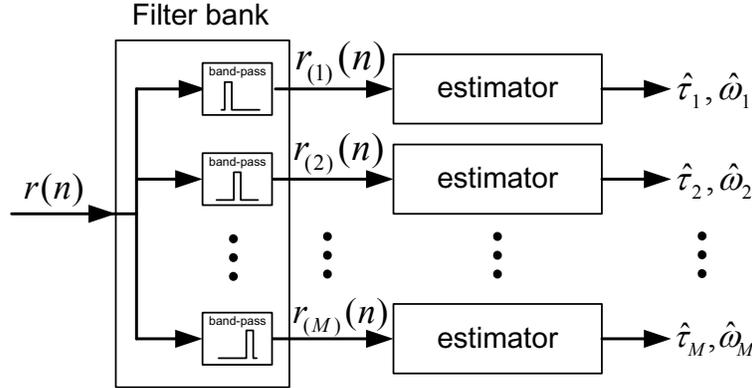


Figure 36: Timing and frequency offsets estimation for subband CAS uplink OFDMA

where $z_{(i)}(n)$ is the signal transmitted from the i th user and can be expressed as

$$z_{(i)}(n) = e^{j2\pi\omega_i n/N} \sum_{l=0}^{L-1} h_i(l)x_i(n-l-\tau_i) \quad (76)$$

and

$$x_i(n) = \frac{1}{\sqrt{N}} \sum_{k \in \mathcal{J}_i} X(k)e^{-j2\pi kn/N} \quad (77)$$

It can be seen from eq. (75) that the received signal at BS is the combination of the signal from all active users. Thus, the uplink synchronization is a multi-parameter estimation procedure. To guarantee that the residual synchronization errors of the uplink transmission are much smaller than that of a completely asynchronous system, before uplink transmission, an initial synchronization is performed during downlink procedure. The timing offsets among users in the uplink are mainly due to different propagation distances between users and BS. The frequency offsets between users and BS are caused by the Doppler spread and/or the instability of local oscillators. Generally, after the synchronization performed via downlink transmission, the CFO can be guaranteed to be in a small range. In this chapter, we assume the frequency offset is smaller than half the subcarrier spacing.

To combat the residual synchronization errors simply and directly, a method based on downlink control channel is suggested in [136, 98], where the synchronization parameters are estimated at the BS and adjustment is performed at the user side based on the information derived from feedback channel. A similar idea is adopted in IEEE 802.16e [137] standard to accomplish the synchronization task. By using interference cancellation or multiuser detection algorithms, e.g., [138, 139, 140, 141] and references therein, the effects or multiple frequency offsets also can be mitigated at the BS at the price of increased receiver complexity. In this chapter, we focus only on the estimation of the timing and frequency offsets at the BS. Dependent on the CAS employed, the synchronization task in uplink OFDMA can be categorised into three cases as explained next.

4.2. Synchronisation with subband CAS

The ML estimation of timing and frequency offsets for subband CAS uplink OFDMA was first studied in [136], where the users' signals are separated by a bank of band-pass filters at BS as shown in Figure 36. After signal separation, the timing and frequency offsets can be estimated independently for each active user, which is similar to the downlink OFDMA case. Since perfect signal separation is impractical, the filtered i th subband signal $r_{(i)}(n)$ can be written as

$$r_{(i)}(n) = z_{(i)}(n) + \Phi_{(i)}(n) + v_{(i)}(n) \quad (78)$$

where the $\Phi_{(i)}(n)$ and $v_{(i)}(n)$ denote the inter-carrier interference due to imperfect separation and noise corresponding to i th subband signal. Possible ways of reducing the interference term $\Phi_{(i)}(n)$ include the adoption of higher order band-pass filters or increasing the number of guard carriers used between adjacent frequency bands. The timing and frequency offsets estimators in [136] exploit the redundancy of CP; the estimates for the i th user can be expressed as:

$$\hat{\tau}_i = \arg \max_{\tilde{\tau}_i} \{ \Lambda_i(\tilde{\tau}_i) - \rho_i C_i(\tilde{\tau}_i) \} \quad (79)$$

$$\hat{\omega}_i = \frac{1}{2\pi} \arg \{ \Lambda_i(\hat{\tau}_i) \} \quad (80)$$

where

$$\Lambda_i(\tilde{\tau}_i) = \sum_{n=\tilde{\tau}_i-N_g}^{\tilde{\tau}_i-1} r_{(i)}^*(n) r_{(i)}(n+N) \quad (81)$$

$$C_i(\tilde{\tau}_i) = \frac{1}{2} \sum_{n=\tilde{\tau}_i-N_g}^{\tilde{\tau}_i-1} (|r_{(i)}(n)|^2 + |r_{(i)}(n+N)|^2) \quad (82)$$

and $\rho_i = \sigma_x^2 / (\sigma_x^2 + \sigma_v^2)$. The estimates in (79) and (80) are one-shot estimators in the sense that the estimates are based on the observation of a single OFDM symbol. More accurate estimates can be obtained by averaging the cost function over Q successive OFDM blocks as

$$\hat{\tau}_i = \arg \max_{\tilde{\tau}_i} \left\{ \sum_{q=0}^{Q-1} (\Lambda_i(\tilde{\tau}_i + q(N + N_g)) - \rho_i C_i(\tilde{\tau}_i + q(N + N_g))) \right\} \quad (83)$$

$$\hat{\omega}_i = \frac{1}{2\pi} \arg \left\{ \sum_{q=0}^{Q-1} \Lambda_i(\hat{\tau}_i + q(N + N_g)) \right\} \quad (84)$$

An alternative blind scheme to obtain estimates of timing and frequency offsets for subband CAS can be found in [142]. As pointed out in [143], the subband CAS offers the possibility of separating signals from different users through a simple filter bank even in a completely asynchronous scenario with arbitrarily large timing errors. Synchronisation algorithms for the downlink OFDMA can be easily extended to subband-based OFDMA systems. On the other hand, grouping subcarriers together makes systems vulnerable to frequency-selective fading. The adoption of an interleaved CAS can provide users with some form of frequency diversity at the expense of slightly increasing the complexity of synchronization.

4.3. Synchronisation with interleaved CAS

Interleaved subcarrier allocation minimises the distances between subcarriers assigned to different users; hence, in the presence of frequency synchronization errors, signals from different users will overlap in the time domain and interfere with each other in the frequency domain due to loss of orthogonality [144]. Thus, it is a challenging task to separate the multiple user signals compared to the subband CAS. However, advanced signal-processing algorithms, e.g., subspace decomposition based methods, can be employed to reduce the synchronization complexity of interleaved systems.

Subspace-based CFO estimation algorithms are studied in [144] and later in [145, 146] for the uplink OFDMA systems; the key is to exploit the periodic structure of the interleaved transmission. Generalised subspace-based CFO estimation also have been studied by [125, 130, 128, 112] for single user OFDM systems. Let $\mathcal{J}_i = \{\eta_i + pK_u\}_{p=0}^{P-1}$ denote the indices set of the subcarriers allocated to i th user, where η_i is an integer in the interval $[0, K_u - 1]$ and $\eta_i \neq \eta_j$ if $i \neq j$, K_u is the maximum number of users and $P = N/K_u$. We assume that the total number of subcarriers N is an integer multiple of K_u in this chapter. The signal from i th user, given in equation (76), can be re-written as

$$z_{(i)}(n) = \alpha e^{j2\pi(\omega_i + \eta_i)n/N} \sum_{p=0}^{P-1} H_i(\eta_i + pK_u) X_i(\eta_i + pK_u) e^{j2\pi p(n - \tau_i)/P} \quad (85)$$

for $n = 0, \dots, N-1$ and $\alpha = e^{-j2\pi\eta_i\tau_i/N}$. It can be found readily that

$$z_{(i)}(n + \mu P) = e^{j\mu\theta_i} z_{(i)}(n) \quad (86)$$

for $n = 0, \dots, P-1$, $\mu = 0, \dots, K_u - 1$ and $\theta_i = 2\pi(\omega_i + \eta_i)/K_u$. From eq. (75), we get

$$r(n + \mu P) = \sum_{i=1}^M e^{j\mu\theta_i} z_{(i)}(n) + v(n + \mu P) \quad (87)$$

$$n = 0, \dots, P-1, \mu = 0, \dots, K_u - 1$$

We arrange the $\{r(n)\}_{n=0}^{N-1}$ samples into a $K_u \times P$ matrix

$$\mathbf{R} = \begin{bmatrix} r(0) & \cdots & r(P-1) \\ r(P) & \cdots & r(2P-1) \\ \vdots & \ddots & \vdots \\ r((K_u-1)P) & \cdots & r(N-1) \end{bmatrix}_{K_u \times P} \quad (88)$$

Letting R_p denote the p th column of matrix \mathbf{R} , we have that

$$R_p = \mathbf{G}z_p + v_p, \quad p = 0, \dots, P-1 \quad (89)$$

where $z_p = [z_{(1)}(p) \ z_{(2)}(p) \ \cdots \ z_{(M)}(p)]^T$ and \mathbf{G} is a $(K_u \times M)$ matrix given by

$$\mathbf{G} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{j\theta_1} & e^{j\theta_2} & \cdots & e^{j\theta_M} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j(K_u-1)\theta_1} & e^{j(K_u-1)\theta_2} & \cdots & e^{j(K_u-1)\theta_M} \end{bmatrix}_{K_u \times M}$$

Once we express the received samples as in eq. (87), CFO estimation can be carried out using a signal subspace decomposition approach. The dimension of the null subspace is dictated by the number of null subcarriers, which is equal to $(K_u - M)P$. The main idea behind the low-complexity CFO estimation algorithms studied in [144, 145, 146] is to estimate the θ_i 's, which are distinct from each other if $|\omega_i| < 0.5$. Indeed, the θ_i 's in eq. (87) cause phase shifts to identical P -sample long segments. Hence, the subspace approach can in a way be seen as an extension of the repetitive-slot CFO estimation approach (see e.g., [111], [108]) to the case of multiple CFO estimation. Employing the estimation of signal parameters via rotational invariance technique (ESPRIT) algorithm, the multiple CFOs can be estimated using the following steps:

Step 1) Arrange the received signal $\{r(n)\}_{n=0}^{N-1}$ into matrix \mathbf{R} .

Step 2) The covariance matrix $\mathbf{\Omega} = \text{E}\{R_p R_p^H\}$ of R_p is estimated by

$$\hat{\mathbf{\Omega}} = \frac{1}{P} \mathbf{R} \mathbf{R}^H$$

Step 3) Compute singular value decomposition (SVD) of $\hat{\mathbf{\Omega}}$:

$$\hat{\mathbf{\Omega}} = [\mathbf{U}_s \ \mathbf{U}_z] \begin{bmatrix} \Sigma_s & \mathbf{0} \\ \mathbf{0} & \Sigma_z \end{bmatrix} \begin{bmatrix} \mathbf{U}_s^H \\ \mathbf{U}_z^H \end{bmatrix} \quad (90)$$

where \mathbf{U}_s is a $K_u \times M$ matrix composed of M eigenvectors corresponding to the M largest eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_M$ and \mathbf{U}_z is a $K_u \times (K_u - M)$ matrix composed of $K_u - M$ eigenvectors corresponding to the rest eigenvalues $\lambda_{M+1} \geq \dots \geq \lambda_{K_u}$.

Step 4) Let \mathbf{U}_{s1} denote the first $(K_u - 1)$ rows of \mathbf{U}_s and \mathbf{U}_{s2} denote the last $(K_u - 1)$ rows of \mathbf{U}_s . The θ_i 's are estimated as

$$\hat{\theta}_i = \angle(\beta_i) \quad (91)$$

where $\angle(\cdot)$ denotes the angle of the complex number and $\{\beta_i\}_{i=0}^{M-1}$ are the eigenvalues of

$$\mathbf{\Xi} = (\mathbf{U}_{s1}^H \mathbf{U}_{s1})^{-1} \mathbf{U}_{s1}^H \mathbf{U}_{s2}$$

Step 5) After estimating the $\hat{\theta}_i$'s, the estimate of CFO ω_i can be computed as

$$\hat{\omega}_i = \frac{K_u \hat{\theta}_i}{2\pi} - \eta_i, \quad i = 1, \dots, M \quad (92)$$

Another subspace-based method, the spectral multiple signal classification (MUSIC) algorithm can also be applied to estimate the multiple CFOs by replacing Step 4 above by

Step 4) Find the M largest peaks of following metric:

$$\Lambda(\tilde{\theta}) = \frac{1}{|a^H(\tilde{\theta}) \mathbf{U}_z \mathbf{U}_z^H a(\tilde{\theta})|^2} \quad (93)$$

where $a(\tilde{\theta}) = [1, e^{j\tilde{\theta}}, \dots, e^{j(K_u-1)\tilde{\theta}}]^T$.

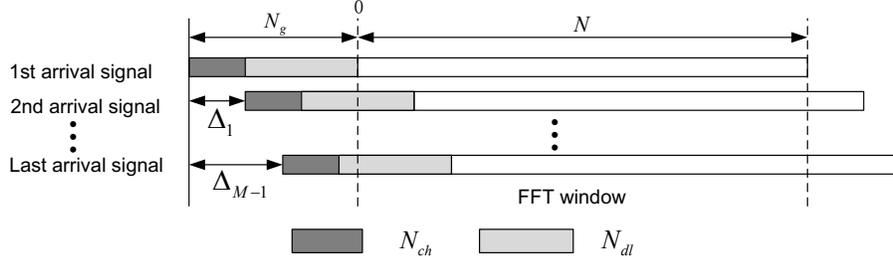


Figure 37: CP for quasi-synchronous uplink OFDMA uplink

As compared in [146], the ESPRIT based estimation algorithm outperforms the Spectral MUSIC based estimation algorithm at low SNR region. Moreover, ESPRIT algorithm avoids the search operation required by Spectral MUSIC. Thus, the ESPRIT based estimation algorithm is preferred in practice. As shown in [146], the estimation error of CFOs can be reduced by using more than one OFDM block.

To apply the subspace-based CFO estimation algorithm, we have to find the starting point of the receive signal first. As argued in [144], the CFO estimation algorithm is applicable to a quasi-synchronous system. The starting point is determined by the downlink procedure and the effect of timing offsets due to propagation delay can be removed by introducing a long CP. As shown in Figure 37, the CP is composed of two parts $N_g = N_{ch} + N_{dl}$, where N_{ch} is the portion of the CP for accommodating channel delay spreads, while the additional N_{dl} samples are intended for accommodating different timing offsets among users. To completely remove the inter-block interference, a necessary condition is that $N_{ch} \geq L$ and $N_{dl} \geq \Delta_{M-1}$, where L is the maximum channel delay spread and Δ_{M-1} is the maximum timing offset among users. In this case, the extra overhead, i.e., N_{dl} , will be increased by an increasing Δ_{M-1} . To reduce the overhead, accurate knowledge of the timing offset of each user is necessary to align all user signals at BS. As shown in [147] and [143], a possible way to estimate the multiple timing offsets is to estimate the timing offsets together with the channel responses. Similar to fine timing estimation in downlink OFDMA, the maximum energy criterion for timing offset estimation can be expressed as

$$\tilde{\tau}_i = \arg \max_{\tilde{\tau}_i} \left\{ \sum_{l=\tilde{\tau}_i}^{\tilde{\tau}_i+L-1} |\hat{h}'_i(l)|^2 \right\} \quad (94)$$

where $\hat{h}'_i(l)$ is the l th entry of \hat{h}'_i and $h'_i = [0_{\tau_i} \ h_i \ 0_{L_{ex}-\tau_i-L}]^T$ is the extended channel vector with length L_{ex} . We can set $L_{ex} = N_g$ for simplicity.

Alternative timing estimation algorithms can be found in [148] and [149]. However, the introduced iterative approaches make the algorithms much more complicated compared to the maximum energy criterion discussed above.

4.4. Synchronisation with generalised CAS

As mentioned before, since there is no rigid constraint between subcarriers and users in generalised CAS, this subcarrier allocation scheme is more flexible than the subband or interleaved CAS. The BS can assign the best subcarriers which are currently available to a

user according to the users' channel condition. Thus, the generalised CAS can improve the systems performance significantly. On the other hand, lack of constraint among subcarriers makes the synchronization task even more challenging than that of interleaved CAS.

The joint ML estimation of timing and frequency offsets for the generalised CAS was first studied by Morelli in [98] and a suboptimal solution was proposed based on repetitively transmitted training symbols. However, it assumes that only one new user enters the network at each time, which may be too strict in practical applications. Alternative ML-based synchronization schemes for the generalised CAS are described in [147] and [150], where iterative alternating projection and space-alternating generalised expectation-maximisation (SAGE) algorithms are employed to reduce the complexity of ML estimation, respectively. Similar to the subspace-based algorithm for interleaved CAS, the iterative-based algorithms studied in [147] and [150] are only applicable to a quasi-synchronous system.

Under the quasi-synchronous assumption as shown in Figure 37, after removing the CP, the IBI free received signal expressed in (75) can be re-written in matrix form as

$$r = \sum_{i=1}^M \Gamma(\omega_i) \mathbf{A}_i h'_i + v \quad (95)$$

or equivalently

$$r = \sum_{i=1}^M \Gamma(\omega_i) \mathbf{D}_i(\tau_i) h_i + v \quad (96)$$

where

$$r = [r(0), \dots, r(N-1)]^T \quad (97)$$

$$\Gamma(\omega_i) = \text{diag}\{1, e^{j2\pi\omega_i/N}, \dots, e^{j2\pi\omega_i(N-1)/N}\} \quad (98)$$

$$[\mathbf{A}_i]_{m,n} = [x_i]_{(m-n) \bmod N}, 0 \leq m \leq N-1, n = 0 \leq n \leq N_g - 1 \quad (99)$$

$$[\mathbf{D}_i(\tau_i)]_{m,n} = [x_i]_{(m-n-\tau_i) \bmod N}, 0 \leq m \leq N-1, n = 0 \leq n \leq L-1 \quad (100)$$

$$h_i = [h_i(0), \dots, h_i(L-1)]^T \quad (101)$$

$$h'_i = [0_{\tau_i \times 1}^T \ h_i^T \ 0_{(N_g-L-\tau_i) \times 1}^T]^T \quad (102)$$

where $[\mathbf{A}_i]_{m,n}$ denotes the (m,n) th entry of matrix \mathbf{A}_i and $[x_i]_m$ represents the m th entry of vector x_i . Rewrite (95) as

$$r = \mathbf{B}(\omega) h' + v \quad (103)$$

where $\mathbf{B}(\omega) = [\Gamma(\omega_1) \mathbf{A}_1 \ \Gamma(\omega_2) \mathbf{A}_2 \ \dots \ \Gamma(\omega_M) \mathbf{A}_M]$ and $h' = [(h'_1)^T \ \dots \ (h'_M)^T]^T$. The log-likelihood function for the frequency offsets ω and extended equivalent channel h' can be expressed as

$$\Lambda(\tilde{\omega}, \tilde{h}') = -N \ln(\pi \sigma_v^2) - \frac{1}{\sigma_v^2} \|r - \mathbf{B}(\tilde{\omega}) \tilde{h}'\|^2 \quad (104)$$

where $\tilde{\omega}$ and \tilde{h}' are trial values of ω and h' respectively. Thus, the joint ML estimates of ω and h' can be obtained as

$$\hat{\omega} = \arg \max_{\tilde{\omega}} \{ \|\Pi_B(\tilde{\omega}) r\|^2 \} \quad (105)$$

$$\hat{h}' = (\mathbf{B}^H(\hat{\omega}) \mathbf{B}(\hat{\omega}))^{-1} \mathbf{B}^H(\hat{\omega}) r \quad (106)$$

where $\Pi_B(\tilde{\omega}) = \mathbf{B}(\tilde{\omega}) (\mathbf{B}^H(\tilde{\omega})\mathbf{B}(\tilde{\omega}))^{-1} \mathbf{B}^H(\tilde{\omega})$. The maximisation in (105) requires a grid-search over the multidimensional domain spanned by $\tilde{\omega}$, which is too cumbersome in practice [147]. A simple way to reduce the complexity is to use the iterative alternating projection method studied in [147].

Let $\hat{\omega}_i^{(k)}$ denote the estimate of ω_i at the k th iteration and define the $M - 1$ dimensional vector $\hat{\omega}_i^{(k)}$ as

$$\hat{\omega}_i^{(k)} = [\hat{\omega}_1^{(k+1)}, \dots, \hat{\omega}_{i-1}^{(k+1)}, \hat{\omega}_{i+1}^{(k)}, \dots, \hat{\omega}_M^{(k)}]^T \quad (107)$$

At the i th step of the $(k + 1)$ th iteration, the estimate of ω_i is updated by the alternating projection frequency estimator (APFE) as

$$\hat{\omega}_i^{(k+1)} = \arg \max_{\tilde{\omega}_i} \left\{ \|\Pi_B(\tilde{\omega}_i, \hat{\omega}_i^{(k)})r\|^2 \right\} \quad (108)$$

Exploiting the structure of $\mathbf{B}(\tilde{\omega}_i, \hat{\omega}_i^{(k)})$, the estimator (108) can be further simplified as

$$\hat{\omega}_i^{(k+1)} = \arg \max_{\tilde{\omega}_i} \left\{ \|\Pi_{C_B}(\tilde{\omega}_i, \hat{\omega}_i^{(k)})r\|^2 \right\} \quad (109)$$

where

$$\Pi_{C_B}(\tilde{\omega}_i, \hat{\omega}_i^{(k)}) = \mathbf{C}_B(\tilde{\omega}_i, \hat{\omega}_i^{(k)}) \left(\mathbf{C}_B^H(\tilde{\omega}_i, \hat{\omega}_i^{(k)})\mathbf{C}_B(\tilde{\omega}_i, \hat{\omega}_i^{(k)}) \right)^{-1} \mathbf{C}_B^H(\tilde{\omega}_i, \hat{\omega}_i^{(k)}) \quad (110)$$

$$\mathbf{C}_B^H(\tilde{\omega}_i, \hat{\omega}_i^{(k)}) = \mathbf{I}_N - \Pi_C(\hat{\omega}_i^{(k)})\Gamma(\tilde{\omega}_i)\mathbf{A}_i \quad (111)$$

$$\Pi_C(\hat{\omega}_i^{(k)}) = \mathbf{C}(\hat{\omega}_i^{(k)}) \left(\mathbf{C}^H(\hat{\omega}_i^{(k)})\mathbf{C}(\hat{\omega}_i^{(k)}) \right)^{-1} \mathbf{C}^H(\hat{\omega}_i^{(k)}) \quad (112)$$

$$\mathbf{C}(\hat{\omega}_i^{(k)}) = \left[\Gamma(\hat{\omega}_1^{(k+1)})\mathbf{A}_1 \cdots \Gamma(\hat{\omega}_{i-1}^{(k+1)})\mathbf{A}_{i-1} \Gamma(\hat{\omega}_{i+1}^{(k)})\mathbf{A}_{i+1} \cdots \Gamma(\hat{\omega}_M^{(k)})\mathbf{A}_M \right] \quad (113)$$

Computing $\Pi_{C_B}(\tilde{\omega}_i, \hat{\omega}_i^{(k)})$ only requires the inversion of a $N_g \times N_g$ matrix, which is significantly less complex than computing $\Pi_B(\tilde{\omega}_i, \hat{\omega}_i^{(k)})$. From (109), we see that the M -D search required by the ML estimator (105) is split into a series of 1-D maximisation problems, and is this much more effective than the original maximisation problem. After obtaining the frequency estimates $\hat{\omega}$, the ML estimates of timing offsets can be derived from eq. (96) as

$$\hat{\tau} = \arg \max_{\tilde{\tau}} \left\{ \|\Pi_\Psi(\hat{\omega}, \tilde{\tau})r\| \right\} \quad (114)$$

where

$$\Pi_\Psi(\hat{\omega}, \tilde{\tau}) = \Psi(\hat{\omega}, \tilde{\tau}) (\Psi^H(\hat{\omega}, \tilde{\tau})\Psi(\hat{\omega}, \tilde{\tau}))^{-1} \Psi^H(\hat{\omega}, \tilde{\tau}) \quad (115)$$

$$\Psi(\hat{\omega}, \tilde{\tau}) = [\Gamma(\hat{\omega}_1)\mathbf{D}_1(\tilde{\tau}_1) \cdots \Gamma(\hat{\omega}_M)\mathbf{D}_M(\tilde{\tau}_M)] \quad (116)$$

Similar to the problem in (105), the maximisation problem in (114) can be efficiently solved by resorting to iterative alternating projection methods; the resulting estimator is referred to as alternating projection timing estimator (APTE) in [147]. Since the timing and frequency estimators introduced above are iterative, initial estimates of ω and τ , referred to as $\hat{\omega}^{(0)}$ and $\hat{\tau}^{(0)}$ respectively, are required. A simple way to initialise the estimates of CFOs is to

use the expected value of ω_i , i.e. $\hat{\omega}_i^{(0)} = 0$. Alternatively, $\hat{\omega}^{(0)}$ can be taken as the output of the frequency estimator proposed in [151]. The initial estimates $\hat{\tau}^{(0)}$ can be obtained by first estimating h'_i according to eq. (106) and then exploiting the specific structure of \hat{h}_i . The index of the first significant element of \hat{h}'_i is taken as $\hat{\tau}_i^{(0)}$.

Another alternative approach to avoid the multidimensional search in ML based estimator is employing the SAGE algorithm as in [149] and [150]. Since the estimates of timing offsets can be obtained via (94), we assume that the timing offsets are estimated first for SAGE algorithm. From eq. (96), the i th cycle of the k th iteration of the SAGE algorithm can be performed as follows.

1. Expectation step: Compute

$$y_i^{(k)} = r - \sum_{j=1}^{i-1} \hat{r}_j^{(k)} - \sum_{j=i+1}^M \hat{r}_j^{(k-1)} \quad (117)$$

2. Maximisation step: The likelihood function for the unknown frequency offset ω_i and channel h_i can be expressed as

$$\Lambda(\tilde{\omega}_i, \tilde{h}_i) = -N \ln(\pi \sigma_v^2) - \frac{1}{\sigma_v^2} \|y_i^{(k)} - \Gamma(\tilde{\omega}_i) \mathbf{D}_i(\tilde{\tau}_i) \tilde{h}_i\|^2 \quad (118)$$

Thus, the joint ML estimates of ω_i and h_i can be written as

$$\hat{\omega}_i^{(k)} = \arg \max_{\tilde{\omega}_i} \left\{ \|\Pi_W(\tilde{\omega}_i) y_i^{(k)}\|^2 \right\} \quad (119)$$

$$\hat{h}_i^{(k)} = (\mathbf{D}_i^H(\hat{\tau}_i) \mathbf{D}_i(\hat{\tau}_i))^{-1} \mathbf{D}_i^H(\hat{\tau}_i) \Gamma^H(\hat{\omega}_i^{(k)}) y_i^{(k)} \quad (120)$$

where $\Pi_W = \Gamma(\tilde{\omega}_i) \mathbf{D}_i(\hat{\tau}_i) (\mathbf{D}_i^H(\hat{\tau}_i) \mathbf{D}_i(\hat{\tau}_i))^{-1} \mathbf{D}_i^H(\hat{\tau}_i) \Gamma^H(\tilde{\omega}_i)$. After obtaining the estimates $\hat{\omega}_i^{(k)}$ and $\hat{h}_i, \hat{r}_i^{(k)}$ which is utilised in the Expectation step of the next cycle or iteration can be updated as

$$\hat{r}_i^{(k)} = \Gamma(\hat{\omega}_i^{(k)}) \mathbf{D}_i(\hat{\tau}_i) \hat{h}_i^{(k)} \quad (121)$$

Again, the initial estimates of CFOs for the SAGE algorithm can be obtained via iterative alternating projection methods. Moreover, we can see from eq. (121) that initial channel estimates are required for the SAGE algorithm while no such requirement is needed for the iterative alternating projection methods. Inaccurate channel estimates will deteriorate the SAGE performance significantly. With the aim of obtaining a tradeoff between performance and complexity, several non-ML based multiple CFOs estimators were proposed in [152] and [153].

4.5. Simulation results for uplink OFDMA systems

The synchronization of uplink OFDMA systems depends on the subcarrier allocation schemes. Here, we compare the synchronization performance of subspace based and ML based methods, based on interleaved CAS. The total number of subcarriers is set to 512 with a CP of

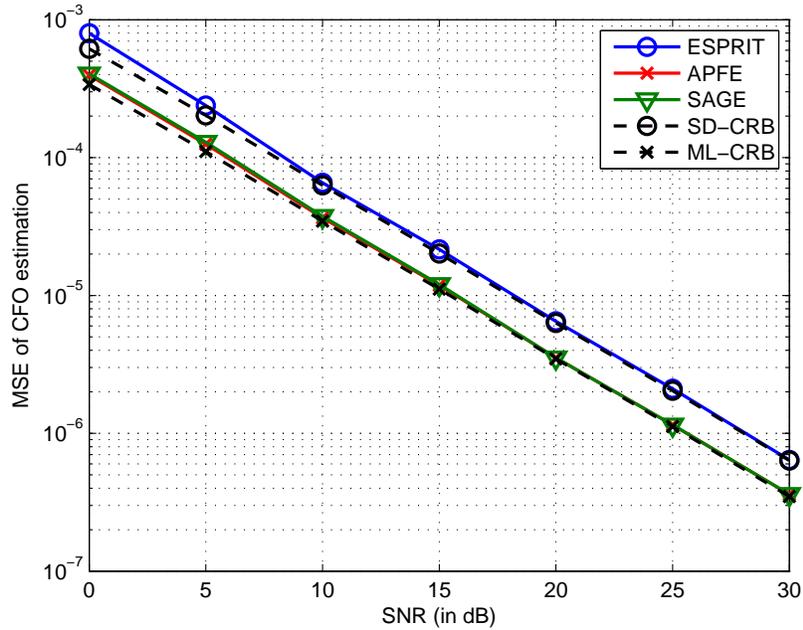


Figure 38: Mean square error of CFO estimates for uplink OFDMA without timing offsets

length 32, i.e., $N = 512$ and $N_g = 32$. The maximum and active numbers of users are set to 4 and 3, respectively, i.e., $K_u = 4$ and $M = 3$. The channel with the exponential power delay profile, introduced previously, is considered. The maximum channel delay spread L is equal to 16 and the channels for different users are assumed uncorrelated. We set the CFOs of the three users to -0.1 , -0.2 and 0.3 , respectively.

Figure 38 shows the CFO estimation performance without timing offsets. From eq. (95), the CRB of ML based CFO estimation algorithm is given in [147] as

$$\text{var}\{\hat{\omega}_i\} \geq \text{CRB}\hat{\omega}_i = \frac{N^2\sigma_v^2}{8\pi^2} \left[\left(\Re \left\{ \Psi^H \Pi_{\mathbf{B}}^\perp \Psi \right\} \right) \right]_{i,i} \quad (122)$$

where $\Psi = [\Psi_1, \dots, \Psi_M]$, $\Psi_i = \mathbf{W}\Gamma(\omega_i)\mathbf{A}_i h_i'$ with $\mathbf{W} = \text{diag}\{0, 1, \dots, N-1\}$; $\Pi_{\mathbf{B}}^\perp = \mathbf{I} - \Pi_{\mathbf{B}}(\omega)$. The CRB of subspace decomposition based CFO estimation algorithm is given in eq. (??). From Figure 38, we can find that ML based algorithms, i.e., APFE and SAGE, have identical performances and provide nearly 3-dB gains over ESPRIT algorithm. This can be ascribed to the fact that subspace decomposition based CFO estimation is actually a non-data aided method. Both ML and subspace decomposition based CFO estimation algorithms can achieve their 'corresponding CRBs'.

Under the quasi-synchronous assumption, we assume that the first arriving signal is known perfectly by the BS and the timing offsets of other two user's signals are normally distributed in an interval $[0, 16]$. Figure 39 shows the timing estimation performance of ESPRIT and SAGE based estimators, and the ML estimator (114), i.e., APTE. As we can see from the figure, the ML estimator outperforms the maximum channel energy based estimators significantly at the price of increased computational complexity. Since APFE and

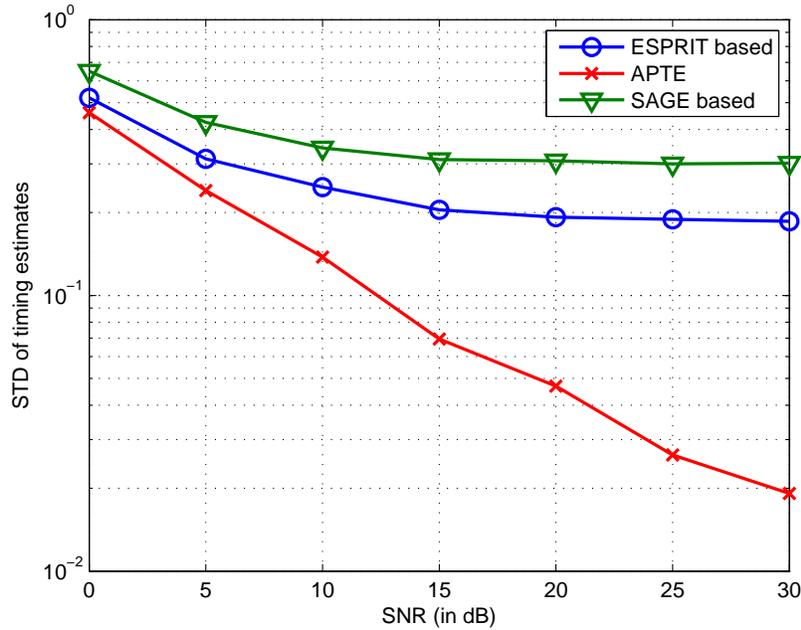


Figure 39: Standard deviation of timing estimates for uplink OFDMA

ESPRIT based CFO estimators can work without timing offset information, we estimate CFO prior to timing offset estimation. On the other hand, timing offsets information is crucial to the SAGE based CFO estimator. Thus we perform timing offset estimation before SAGE based CFO estimation. From Figure 39, we see that ESPRIT based timing estimator outperforms SAGE. This can be ascribed to the fact that employing the CFO estimates obtained in the ESPRIT based CFO estimator improves timing estimation. Figure 40 shows the CFO estimation performance with timing offsets estimation. We see that the performance of SAGE based CFO estimator is affected significantly by the inaccuracy of timing estimation. On the other hand, both ESPRIT and APFE estimators are robust to timing inaccuracies.

5. Network synchronization

In this section, we discuss the challenges in the network time synchronization problem and the performance metrics of interest. We provide a brief overview of current devices and their limitations, and describe common clock models. We then describe a taxonomy of network time sync protocols and provide some examples.

As stated in the Introduction to this chapter, network synchronization is a well-studied topic with an extensive history, especially for wired networks, e.g., see Lindsey et al. [1], Bregni [2], and references therein. Typically, these works assumed high quality devices, availability of fine control of the network, extensive connectivity with little or no mutual interference, as well as often assuming known (or repeatable and measurable) propagation

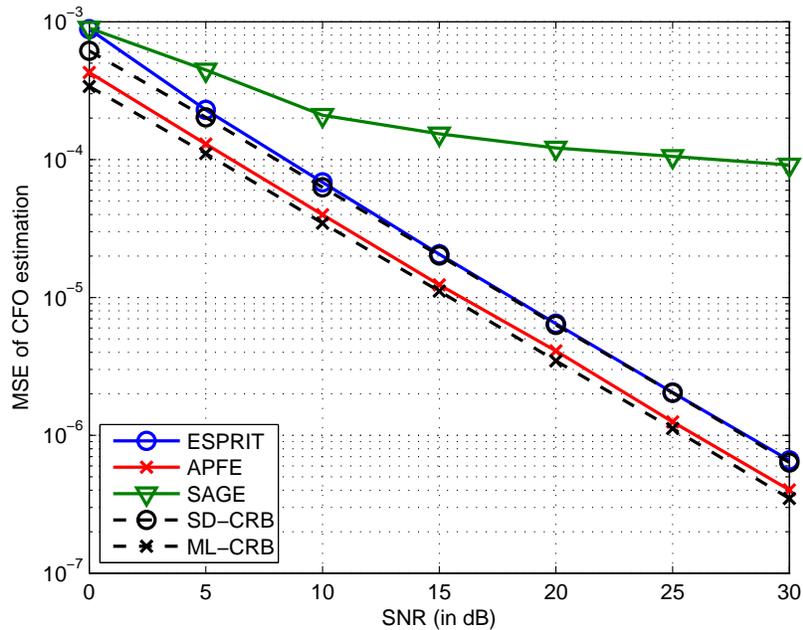


Figure 40: Mean square error of CFO estimates for uplink OFDMA with timing offsets

and processing delays [3]. Surveys of WSN sync protocols may be found in the papers by Sivrikaya and Yener [4], Johannessen [5], Sundararaman et al. [6], [3], and [7].

The network synchronization problem is to ensure that all nodes in the network operate on a common clock, i.e., have a common time reference.

Challenges in synchronization in WSN stem from several sources, broadly related to the transmitter, the propagation channel, and the receiver.

1. Channel conditions (such as fading, shadowing, interference) lead to time-varying connectivity even for static nodes: scatterers move in any case; mobility adds its own challenges.
2. The devices are cheap and clocks drift, often erratically, due to fluctuations in ambient temperature, and with age. The time difference between two clocks may be fixed (a fixed offset) or may vary with time (due to clock oscillator frequency drifts).
3. Queuing and processing delays are variable (thus rendering it difficult to use standard protocols such as NTP). There is variability in the time it takes for a packet to go from the application layer to the MAC layer, variable delays within the MAC layer (the major source of error), in packet generation and transmission at the PHY layer. There are similar variations at the receiver, including inaccuracy in detecting packet arrival.
4. Variabilities in propagation time due to non-line of sight issues, and non-reciprocity of the channel. Typically, the propagation time is negligible compared with the queuing and processing delays.

5. Communication rates are variable in a large network; if a node is involved infrequently in regular communications, then ‘heartbeat’ signals may be essential to keep the node in quasi-sync with the rest of the network, and thus connected to the network.
6. Protocols must be scalable to a large number of nodes, and must deal with heterogeneity of nodes.
7. These devices are often battery powered so that energy is a finite resource, and energy consumption directly affects node lifetime.
8. Given energy constraints, WSN nodes must exploit external assets which may have more relaxed energy constraints; e.g., basestations, UAVs, various broadcast beacons.

Metrics: How should one evaluate the performance of a network sync protocol? When GPS is available, a reasonable metric could be the bias compared to the ‘true’ time. But access to GPS can be difficult, particularly indoors, under canopy, and in other challenging conditions. Often, it suffices that the nodes converge on some common time reference (regardless of whether it is ‘true’). Some commonly used metrics are:

- *Synchronization accuracy:* Worst case (or average case) pair wise error between any one-hop neighbors
- *Energy efficiency:* The number of packet transmissions and receptions necessary to achieve sync, and the rate and frequency of messages that need to be exchanged to maintain sync.
- *Synchronization convergence time:* The time taken for all nodes (or a given percentage of nodes) to be in sync with one-hop neighbors.
- *Fault tolerance:* The robustness of sync schemes under (intermittent) failure of (critical) nodes and/or links; robustness to (slow) time variations in clock parameters and clock jitter.
- *Scalability with network size:* Does the sync-error increase with size? Does convergence time increase (only) with the diameter of the network? Or other aspects of topology such as degree distribution?
- *Impact of stochastic channel conditions:* How well does the protocol perform in the presence of stochastic channel conditions (congestion, mobility, duty cycling, queueing delays, propagation delays, processing times)?
- *Engineering desing:* Is the protocol simple vs. complex

From the above questions, it is clear that a given protocol offers a set of alternatives in this rich tradespace.

5.1. Clock Models

Let $T_k(t)$ denote the local time at node k , where t denotes the ‘true’ time; often, we will drop the node index k . Formally, *clock drift* is defined as

$$\rho(t) = \frac{dT(t)}{dt} - 1.$$

A reasonable assumption is that the drifts are bounded, and the clocks do not run backwards, which translate to

$$|\rho(t)| < \rho_{\max}; \quad \rho(t) > -1.$$

A Taylor expansion of the local clock time $T(t)$ wrt the global clock t yields

$$T_k(t) = \alpha_k + \beta_k t + \gamma_k t^2 + \dots \quad (123)$$

wherein α is the *offset* and β the *skew*. The quadratic term, denoted by γ , is typically used only to test for departures from the linear model. Skew has been modeled as an AR process in [154]

	Accuracy (PPM)	Power	Lifetime in hours AA battery	
GPS	$10^{-8} \sim$	10^{-11}	180 mW	16.7 hours
Chip-Scale Atomic Clock	10^{-11}	30 mW	100 hours	
MCXO	3×10^{-8}	75 mW	40 hours	
TCXO	6×10^{-6}	6 mW	21 days	
Watch clock	200×10^{-6}	$1 \mu\text{W}$	342 years	

Table 2: Compariosn of clock characteristics

5.2. Net Sync Protocols

With the above background, we can broadly classify net sync protocols into four broad categories:

1. Broadcast protocols: Based on the notion of broadcast, possibly over a hierarchical tree topology
2. Distributed synchronization: builds consensus on clock parameters in a peer-to-peer setting.
3. Unilateral sync to an external (broadcast) reference clock.

Other classifications are possible depending upon the viewpoint: e.g., client (sensor node) initiated vs. server (gateway) initiated.

Many WSN synchronization protocols have been proposed. Popular ones include the Reference Broadcast Systems (RBS) [155], a time-stamp transformation approach based on bounded offsets [156], the Tiny/Mini-Sync (TMS) protocol [157], Timing-sync Protocol for Sensor Networks (TPSN) [158], Lightweight Tree-based Synchronization (LTS) [159], and the networked Control Time Protocol (CTP) [160]. Probabilistic approaches were considered in [161], and refined and extended in [162]. Bounds under various assumptions are derived in [163], [164] and [165].

Among the above protocols, several primary themes emerge. One natural and common notion is the use of time-stamps: time-stamp a packet with the transmitter's clock, time-stamp the reception time, use these stamps to estimate the round trip time (RTT), which is then used to synchronize the two clocks. RTT is often highly variable, and often has a heavy-tailed distribution, which naturally calls for the use of robust estimation techniques. Reliability increases as the number of such exchanges increases, but with a concomitant increase in delay and energy, and sometimes (more than linear) complexity. A second recurring theme is that the estimation of relative clock offset and skew can be cast as a linear estimation problem. and complexity.

Unilateral synchronization

Assume that an external source broadcasts time-stamped messages at 'true' times $T(i)$, which are received by a node at $R(i)$ on its local clock. Then from the clock model considered earlier in (123), we have

$$R(i) = \alpha + \beta T(i) + \varepsilon(i), \quad i = 1, \dots, I \quad (124)$$

where n is the number of observations. Here $\varepsilon(i)$ represents the modeling errors. Let $\mathbf{1}$ be $I \times 1$ vector of ones, $\mathbf{t} = [T(1), \dots, T(I)]'$, $\mathbf{r} = [R(1), \dots, R(I)]'$, $\mathbf{Z} = [\mathbf{1}, \mathbf{t}]$. Then linear regression yields

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = (\mathbf{Z}'\mathbf{Z}^{-1}\mathbf{Z}'\mathbf{r})$$

which is the best linear estimate if the $\varepsilon(i)$ are zero-mean. Note that in this unilateral scheme, one cannot account separately for the propagation delay, or a non-zero-mean ε : both are absorbed into the offset. Thus clocks that are at relatively different distances from the broadcast source will not be in sync with each other.

Pairwise synchronization

As in the classical Network Time Protocol (NTP) [166], clock offset can be estimated by exchanging time-stamped messages and computing the round trip time (RTT). If the queuing delays are exponentially distributed with the same mean delay then the MLE of the offset is given by the minimum of the observed delays [167]; if the mean delays vary from node to node, then the bootstrap-bias correction method of [168] can be used. Further details may be found in [169].

In these pair-wise protocols, a node 'B' synchronizes with a node 'A' which is treated as the reference node. Let α and β denote the relative offset and skew of node B wrt node

A. Nodes exchange multiple time-stamped messages as follows. Node A sends a time-stamped message to node B at its local time $T_{A,k}$, where k is the round number. Node B receives it at its local time $R_{B,k}$. At $T_{B,k}$, it sends a time-stamped message which includes $R_{B,k}$; this is received by node A at $R_{A,k}$. K such rounds of message exchange take place. Let δ denote the fixed portion of the delay and $\varepsilon_{AB,k}, \varepsilon_{BA,k}$, the variable portion; delay here includes propagation delay, processing delay and queuing delay. The time-stamps are related via:

$$R_{B,k} = (T_{A,k} + \delta + \varepsilon_{AB,k})\beta + \alpha \quad (125)$$

$$T_{B,k} = (R_{A,k} - \delta - \varepsilon_{BA,k})\beta + \alpha \quad (126)$$

Assuming that the delays are independent and exponentially distributed, it is shown in [167, 169] that the MLE of α is given by

$$\hat{\alpha} = \frac{1}{2} \left[\min_k (R_{B,k} - T_{A,k}) - \min_k (R_{A,k} - T_{B,k}) \right].$$

Let $i = \arg \min_k (R_{A,k} - T_{A,k})$ and $j = \arg \min_{k,k \neq i} (R_{A,k} - T_{A,k})$, which are the first two order statistics of $R_{A,k} - T_{A,k}$. Then the proposed estimator of the skew in [169] is

$$\hat{\beta} = \frac{R_{B,i} + T_{B,i} - R_{B,j} - T_{B,j}}{R_{A,i} + T_{A,i} - R_{A,j} - T_{A,j}}.$$

Once α, β have been estimated, it is easy to estimate the propagation delay as well. These algorithms have been shown to be robust to other delay models.

In the RBS protocol [155], a beacon node transmits a reference packet (as above), but the K receiving nodes exchange time-of-receipt to estimate offsets and skews. Consider (124), now indexed by the receiving node's id:

$$R_i(n) = \alpha_i + \beta_i T(n) + \varepsilon_i(n), \quad (127)$$

Defining $\Delta_{ij} := \alpha_i - \beta_{ij}\alpha_j$ and $\beta_{ij} := \frac{\beta_i}{\beta_j}$, as the relative offset and skew, one can eliminate $T(n)$ to obtain

$$R_i(n) = \beta_{ij}R_j(n) + \Delta_{ij} + \varepsilon_{ij}(n)$$

Nodes i and j can estimate the relative skew and offset via linear regression.

The tree-based sync protocol in [158] is similar to the above algorithms in the estimation part. Here a root node broadcasts a beacon. Its set of 1-hop neighbors (i.e., those who hear the root directly) are called level 1 nodes. Level 1 nodes, in turn, relay the beacon to level-2 nodes, and so on. Nodes at level i synchronize to a parent node at level $i - 1$. The relative skew is assumed to be unity, so the focus is on estimating offsets. Consider a pair of parent-child nodes; one can write the packet reception times as

$$R_{B,k} = T_{A,k} + \alpha + \delta + \varepsilon_{AB,k}$$

$$R_{A,k} = T_{B,k} - \alpha + \delta + \varepsilon_{BA,k}$$

After K messages, one can estimate

$$\begin{aligned}\hat{\alpha} &= \frac{1}{2K} \sum_{k=1}^K (R_{B,k} - R_{A,k} + T_{B,k} - T_{A,k}) \\ \hat{\delta} &= \frac{1}{2K} \sum_{k=1}^K (R_{B,k} + R_{A,k} - T_{B,k} - T_{A,k})\end{aligned}$$

Under the assumption that the noise terms, the ε 's are zero-mean and independent, one obtains the following equation for the variance of the estimators:

$$\text{var}(\hat{\beta}) = \text{var}(\hat{\delta}) = \frac{\sigma_{12}^2 + \sigma_{21}^2}{K}.$$

The TPSN protocol is easy to implement. However, it assumes unit skew. Tine-Sync and Mini-Sync are variations that can cope with skew [157].

Another variation was proposed in [3]. The basic idea is that if node A transmits a pair of time-stamped messages ρ units apart on its own clock, then node B should be able to estimate the relative skew.

5.3. Distributed Clock Sync

The notion of distributed consensus, i.e., distributed agreement on a parameter (vector) via repeated exchange of messages has recently become popular, and several protocols have been proposed. Here, the idea is that there is no single root node and hence no single point of failure. These protocols seek to achieve global consensus via local exchange of information. They rely upon the broadcast nature of the wireless link.

Bio-Inspired Approaches The seminal work of Mirollo and Strogatz [170] introduced the basic idea that a population of “integrate-and-fire” oscillators will start firing simultaneously after a finite time starting from all almost any initial condition. Empirical evidence for this is the well-cited firefly sync [9] and circadian sync [10]. The work was extended to multi-hop and time-varying topologies in [171] who made use of results from Algebraic Graph theory. These results have been recently refined, corrected and extended by [172]. The Mirollo-Strogatz model has been exploited in [173] for network time sync. Another consensus-based approach is the so-called diffusion-based approach of [174].

Consensus-type approaches We consider next one example of a consensus-based approach [175]. This scheme assumes the presence of a reference node (i.e., a root node whose clock is assumed to be correct, and which does not update its clock). Multiple consistent reference nodes may be present as well. For simplicity assume that all nodes have unit skew, and node i has offset α_i wrt the reference node. Using one of the pairwise schemes described earlier, nodes can obtain an estimate of the relative offset

$$y_{i,j} = \alpha_i - \alpha_j + \varepsilon_{i,j}$$

where $\varepsilon_{i,j}$ is measurement noise which we model as zero-mean and with variance $\sigma_{i,j}^2$. A node i would have such an estimate for each $j \in \mathcal{N}_i$, the set of neighbors that it can hear.

At every iteration, a node receives the current estimates of its neighbors which are used to update its own estimate. Node i 's estimate of its offset α_i in the k -th iteration is given by

$$\hat{\alpha}_i(k) = \left(\sum_{j \in \mathcal{N}_i} \frac{1}{\sigma_{i,j}^2} \right)^{-1} \sum_{j \in \mathcal{N}_i} \frac{1}{\sigma_{i,j}^2} (\hat{\alpha}_j(k-1) + y_{ij})$$

The convergence of this algorithm was studied in [175].

Convergence depends upon the topology, the coupling and the measurement statistics. We can associate a graph $G = (V, E)$ with this problem; here V is the set of nodes, and E the set of edges representing one-hop links. The energy (convergence time) spent for achieving consensus is proportional to $1/\lambda_2(L)$ where L is the graph Laplacian and $\lambda_2(\cdot)$ is the second-largest eigenvalue. The algebraic network connectivity (the so-called Fiedler value), $\lambda_2(L)$ can be increased by increasing the transmit power, i.e., making the graph more connected. However, this also increases the total energy consumption, since the total power consumed by the network is proportional to $p_T/\lambda_2(L(p_T))$, where $L(p_T)$ is the Laplacian of the graph corresponding to transmit power p_T . A question then is: does a global trade-off exist between local transmit power, convergence rate and network topology? Analysis in [176] indicates that when the path loss is high, the optimal topology tends to be sparse, with few connections.

6. Conclusions

Bibliography

- [1] W. C. H. K. D. W. C. Lindsey, F. Ghazvinian, “Network synchronization,” *Proceedings of the IEEE*, vol. 73, no. 10, pp. 1445–1467, Oct. 1985.
- [2] S. Bregni, “A historical perspective on telecommunications network synchronization,” *IEEE Communications Magazine*, pp. 150–166, Jun. 1988.
- [3] B. Sadler and A. Swami, “Synchronization in sensor networks: an overview,” in *IEEE MILCOM, 2006*, 2006.
- [4] F. Sivrikaya and B. Yener, “Time synchronization in sensor networks: a survey,” *IEEE Network*, pp. 45–50, Jul. 2004.
- [5] S. Johannessen, “Time synchronization in a local area network,” *IEEE Control Systems Magazine*, pp. 61–69, Apr. 2004.
- [6] A. K. B. Sundararaman, U. Buy, “Clock synchronization for wireless sensor networks: a survey,” *Ad Hoc Networks (Elsevier)*, pp. 281–323, 2005.
- [7] Y. R. Faizulkhakov, “Time synchronization methods for wireless sensor networks: A survey,” *Programming and Computer Software (SpringerLink)*, vol. 33, no. 4, pp. 214–226, 2007.
- [8] P. Galison, *Einstein’s clocks, Poincare’s maps: empires of time*. Norton, 2003.
- [9] S. Strogatz, *SYNC: the emerging science of spontaneous order*. Hyperion, 2003.
- [10] F. D. I. Y.Q. Wang, “On influences of global and local cues on the rate of synchronization of oscillator networks,” *Automatica*, vol. 47, no. 6, Jun. 2011.
- [11] J. Proakis, *Digital Communications*. Mc-Graw-Hill, 2001.
- [12] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [13] S. Kay, *Fundamentals of Statistical Signal Processing, Volume 1: Estimation Theory*. Prentice Hall, 1993.
- [14] E. Tozer, *Broadcast engineer’s Reference Book*. Focal Press, 2004.
- [15] S. Savory, “Digital filters for coherent optical receivers,” *Optic Express*, vol. 16, no. 2, pp. 804–817, 2008.

-
- [16] P. Ciblat, P. Loubaton, E. Serpedin, and G. Giannakis, "Asymptotic analysis of blind cyclic correlation based symbol rate estimation," *IEEE Transactions on Information Theory*, vol. 48, no. 7, pp. 1922–1934, July 2002.
- [17] U. Mengali and A. d'Andrea, *Synchronization Techniques for Digital Receivers*. Plenum Press, 1997.
- [18] H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital Communications Receivers: synchronization, channel estimation, and Signal Processing*. J. Wiley & Sons, 1998.
- [19] C. Georghiades and E. Serpedin, *The Handbook of Communications*. CRC Press, 2002, ch. Synchronization.
- [20] B. Porat, *Digital Signal Processing of Random Signals*. Prentice Hall, 1994.
- [21] A. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Transactions on Information Theory*, 1967.
- [22] G. Forney, "The viterbi algorithm," *Proceedings of the IEEE*, 1973.
- [23] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, April 2003.
- [24] H. V. Trees, *Detection, Estimation, and Modulation Theory*. J. Wiley & Sons, 2003.
- [25] S. Crozier, D. Falconer, and S. Mahmoud, "Least sum of squared errors channel estimation," *IEE Proceedings-F (Radar and Signal Processing)*, vol. 138, no. 4, pp. 371–378, Aug. 1991.
- [26] C. Tellambura, M. Parker, Y. Guo, S. Sheperd, and S. Barton, "Optimal sequences for channel estimation using discrete Fourier transform techniques," *IEEE Transactions on Communications*, vol. 47, no. 2, pp. 230–237, Feb. 1999.
- [27] A. Benveniste, M. Métivier, and P. Priouret, *Adaptive Algorithms and Stochastic Approximations*. Springer-Verlag, 1990.
- [28] C. Herzet, N. Noels, V. Lottici, H. Wymeersch, M. Luise, M. Moeneclaey, and L. Vandendorpe, "Code-aided turbo synchronization," *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1255–1271, June 2007.
- [29] W. Lindsey and M. Simon, *Telecommunications Systems Engineering*. Englewood Cliffs, Prentice Hall, 1972.
- [30] F. Gardner, *Demodulator Reference recovery techniques suitable for digital implementation*. ESTEC contrat no. 6847/86/NL/DG, European Space Agency, 1988.
- [31] K. Mueller and M. Mueller, "Timing recovery in digital synchronous data receivers," *IEEE Transactions on Communications*, vol. 24, no. 5, pp. 516–531, May 1976.

-
- [32] M. Morelli and U. Mengali, "Carrier frequency estimation for transmissions over selective channels," *IEEE Transactions on Communications*, vol. 48, no. 9, pp. 1580–1589, Sep. 2000.
- [33] P. Ciblat and L. Vandendorpe, "On the maximum-likelihood based data-aided frequency offset and channel estimates," in *EURASIP European Signal Processing Conference (EUSIPCO)*, vol. 1, Toulouse (France), Sep. 2002, pp. 627–630.
- [34] D. Rife and R. Boorstyn, "Single-tone parameter estimation from discrete-time observations," *IEEE Trans. on Information Theory*, vol. 20, no. 5, pp. 591–598, Sep. 1974.
- [35] R. M. Aulay and L. Seidman, "A useful form of Barankin lower bound and its application to PPM threshold analysis," *IEEE Trans. on Information Theory*, vol. 15, no. 2, pp. 273–279, Mar. 1969.
- [36] L. Atallah, J.-P. Barbot, and P. Larzabal, "SNR threshold indicators in data-aided phase and frequency synchronization for QAM," *IEEE Signal Processing Letters*, vol. 11, no. 8, pp. 652–654, Nov. 2004.
- [37] P. Ciblat and M. Ghogho, "Blind NLLS carrier frequency-offset estimation for QAM, PSK, and PAM modulations : performance at low SNR," *IEEE Transactions on Communications*, vol. 54, no. 10, pp. 1725–1730, Oct. 2006.
- [38] E. Barankin, "Locally best unbiased estimates," *Annals of Mathematical Statistics*, vol. 20, pp. 447–501, 1949.
- [39] L. Knockaert, "The Barankin bound and threshold behavior in frequency estimation," *IEEE Trans. on Signal Processing*, vol. 45, pp. 2398–2401, Sep. 1997.
- [40] R. M. Aulay and E. Hofstetter, "Barankin bounds on parameter estimation," *IEEE Trans. on Information Theory*, vol. 17, no. 6, pp. 669–676, Nov. 1971.
- [41] I. Reuven and H. Messer, "A Barankin-type lower bound on the estimation error of hybrid parameter vector," *IEEE Trans. on Information Theory*, vol. 43, no. 3, pp. 1084–1093, May 1997.
- [42] —, "On the effect of nuisance parameters on the threshold SNR value of the Barankin bound," *IEEE Trans. on Signal Processing*, vol. 47, pp. 523–527, Feb. 1999.
- [43] P. Forster and P. Larzabal, "On lower bounds for deterministic parameter estimation," in *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 2, Orlando (FL), May 2002, pp. 1137–1140.
- [44] J. Ziv and M. Zakai, "Some lower bounds on Signal Processing estimation," *IEEE Trans. on Information Theory*, vol. 15, no. 3, pp. 386–391, May 1969.
- [45] K. Bell, Y. Steinberg, Y. Ephraim, and H. V. Trees, "Extended Ziv-Zakai lower bound for vector parameter estimation," *IEEE Trans. on Information Theory*, vol. 43, no. 2, pp. 624–637, Mar. 1997.

-
- [46] A. Weiss and E. Weinstein, "Fundamental limitations in passive time-delay estimation – Part I: Narrow-band systems," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 31, no. 2, pp. 472–486, Apr. 1983.
- [47] P. Ciblat and M. Ghogho, "Ziv-zakai bound for harmonic retrieval in multiplicative and additive gaussian noise," in *IEEE Workshop on Statistical Signal Processing (SSP)*, Bordeaux (France), Jul. 2005.
- [48] C. Georghiades and D. Snyder, "The expectation-maximization algorithm for symbol unsynchronized sequence detection," *IEEE Transactions on Communications*, vol. 39, no. 1, pp. 54–61, Jan. 1991.
- [49] E. Panagirci and C. Georghiades, "Joint ml timing and phase estimation in ofdm systems using the em algorithm," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 3, Jul. 2000, pp. 2949 – 2952.
- [50] M. Moeneclaey and G. de Jonghe, "ML-oriented nda carrier synchronization for general rotationally symmetric signal constellations," *IEEE Transactions on Communications*, vol. 42, no. 8, pp. 2531–2533, Aug. 1994.
- [51] M. Morelli and U. Mengali, "Feedforward frequency estimation for PSK : a tutorial review," *European Transactions on Telecommunications*, vol. 9, no. 2, pp. 103–115, March 1998.
- [52] J. Costas, "Synchronous communications," *Proceedings of the IRE*, vol. 44, no. 12, pp. 1713–1718, Dec. 1956.
- [53] D. Taylor, "Introduction tp "synchronous communications"," *Proceedings of the IEEE*, vol. 90, no. 8, pp. 1459–1460, Aug. 2002.
- [54] W. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Elsevier Signal Processing*, vol. 86, no. 4, pp. 639–697, April 2006.
- [55] G. B. Giannakis, F. Panduru, E. Serpedin, and I. Sari, "A bibliography on cyclostationary signal processing," *Elsevier Signal Processing*, vol. 85, no. 12, pp. 2233–2303, Dec. 2005.
- [56] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. on Signal Processing*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.
- [57] F. Gini and G. Giannakis, "Frequency offset and symbol timing recovery in flat-fading channels : a cyclostationary approach," *IEEE Trans. on Communications*, vol. 46, no. 3, pp. 400–411, Mar. 1998.
- [58] A. Dandawate and G. Giannakis, "Asymptotic theory of mixed time averages and k th-order cyclic moment and cumulant statistics," *IEEE Trans. on Information Theory*, vol. 41, no. 1, pp. 216–232, Jan. 1995.
- [59] G. Zhou and G. Giannakis, "Harmonics in multiplicative and additive noise : performance analysis of cyclic estimators," *IEEE Trans. on Signal Processing*, vol. 43, no. 6, pp. 1445–1460, June 1995.

-
- [60] M. Oerder and H. Meyr, "Digital filter and square timing recovery," *IEEE Transactions on Communications*, vol. 36, no. 5, pp. 605–612, May 1988.
- [61] K. Scott and E. Olsz, "Simultaneous clock phase and frequency offset estimation," *IEEE Transactions on Communications*, vol. 43, no. 7, pp. 2263–2270, Jul. 1995.
- [62] M. Ghogho, A. Swami, and T. Durrani, "On blind carrier recovery in time-selective fading channel," in *Asilomar Conference on Signals, Systems and Computers*, vol. 1, 1999, pp. 243–247.
- [63] Y. Wang, E. Serpedin, and P. Ciblat, "Blind feedforward cyclostationarity-based timing estimation for linear modulations," *IEEE Transactions on Wireless Communications*, vol. 3, no. 3, pp. 709–715, May 2004.
- [64] J. Delmas, "Asymptotic performance of second-order algorithms," *IEEE Transactions on Signal Processing*, vol. 50, no. 1, pp. 49–57, Jan. 2002.
- [65] B. Ottersten, P. Stoica, and R. Roy, "Covariance matching estimation techniques for array signal processing applications," *Digital Signal Processing*, vol. 8, pp. 185–210, 1998.
- [66] Y. Wang, P. Ciblat, E. Serpedin, and P. Loubaton, "Performance analysis of a class of non-data aided carrier frequency offset and symbol timing delay estimators for flat-fading channels," *IEEE Transactions on Signal Processing*, vol. 50, no. 9, pp. 2295–2305, Sep. 2002.
- [67] E. Hannan, "Non linear time series regression," *Journal of Applied Probability*, vol. 8, pp. 767–780, 1972.
- [68] ———, "The estimation of frequency," *Journal of Applied Probability*, vol. 10, pp. 510–519, 1973.
- [69] T. Hasan, "Non linear time series regression for a class of amplitude modulated cosinusoids," *Journal of Time Series Analysis*, vol. 3, no. 2, pp. 109–122, Jan. 1982.
- [70] P. Ciblat, P. Loubaton, E. Serpedin, and G. Giannakis, "Performance analysis of blind carrier frequency offset estimators for non-circular transmissions through frequency-selective channels," *IEEE Transactions on Signal Processing*, vol. 50, no. 1, pp. 130–140, Jan. 2002.
- [71] O. Besson and P. Stoica, "Nonlinear least-squares approach to frequency estimation and detection for sinusoidal signals with arbitrary envelope," *Digital Signal Processing*, vol. 9, no. 1, pp. 45–56, Jan. 1999.
- [72] A. Viterbi and A. Viterbi, "Non-linear estimation of PSK-modulated carrier phase with application to burst digital transmissions," *IEEE Trans. on Information Theory*, vol. 29, pp. 543–551, July 1983.
- [73] Y. Wang, E. Serpedin, and P. Ciblat, "Optimal blind nonlinear least-squares carrier phase and frequency offset estimation for general qam modulations," *IEEE Transactions on Wireless Communications*, vol. 2, pp. 1040–1054, Sep. 2003.

-
- [74] S. Colonnese, S. Rinauro, G. Panci, and G. Scarano, "Gain-control-free blind carrier frequency offset acquisition for qam constellations," *IEEE Transactions on Signal Processing*, vol. 58, pp. 349–361, Jan. 2010.
- [75] E. Serpedin, P. Ciblat, G. B. Giannakis, and P. Loubaton, "Performance analysis of blind carrier phase estimators for general qam constellations," *IEEE Transactions on Signal Processing*, vol. 49, pp. 1816–1823, Aug. 2001.
- [76] G. Vazquez and J. Riba, *Signal Processing Advances for Wireless Mobile Communications*. Englewood Cliffs, Prentice-Hall, 2000, vol. 2, ch. 9 (Non-Data-Aided Digital Synchronization).
- [77] A. d'Andrea, U. Mengali, and R. Reggiannini, "The modified Cramer-Rao bound and its application to synchronization problems," *IEEE Transactions on Communications*, vol. 42, no. 2/3/4, pp. 1391–1399, 1994.
- [78] G. Zhou and G. Giannakis, "Harmonics in gaussian multiplicative and additive white noise : Cramer-rao bounds," *IEEE Transactions on Signal Processing*, vol. 43, pp. 1217–1231, May 1996.
- [79] M. Ghogho, A. Swami, and T. Durrani, "Frequency estimation in the presence of Doppler spread : performance analysis," *IEEE Transactions on Signal Processing*, vol. 49, no. 4, pp. 777–789, Apr. 2001.
- [80] F. Gini, R. Reggiannini, and U. Mengali, "The modified Cramer-Rao Bound in vector parameter estimation," *IEEE Transactions on Communications*, vol. 46, no. 1, pp. 52–60, Jan. 1998.
- [81] F. Gini, M. Luise, and R. Reggiannini, "Cramer-rao bounds in the parametric estimation of fading radiotransmission channels," *IEEE Transactions on Communications*, vol. 46, pp. 1390–1398, Oct. 1998.
- [82] M. Moeneclaey, "On the true and the modified Cramer-Rao Bounds for the estimation of a scalar parameter in the presence of nuisance parameters," *IEEE Transactions on Communications*, vol. 46, no. 11, pp. 1536–1544, Nov. 1998.
- [83] F. Rice, B. Cowley, B. Moran, and M. Rice, "Cramer-rao lower bound for QAM phase and frequency estimation," *IEEE Trans. on Communications*, vol. 49, pp. 1582–1591, Sep. 2001.
- [84] H. Steendam and M. Moeneclaey, "Low-SNR limit of the Cramer-Rao Bound for estimating the carrier phase and frequency of a PAM, PSK, or QAM waveform," *IEEE Com. Letters*, vol. 5, no. 5, pp. 218–220, May 2001.
- [85] P. Ciblat, M. Ghogho, P. Larzabal, and P. Forster, "Harmonic retrieval in the presence of non-circular gaussian multiplicative noise : Performance bounds," *EURASIP Signal Processing*, vol. 85, pp. 737–749, April 2005.

-
- [86] N. Noels, H. Steendam, and M. Moeneclaey, "The true cramer-rao bound for carrier frequency estimation from a psk signal," *IEEE Transactions on Communications*, vol. 52, pp. 834–844, May 2004.
- [87] P. Stoica and B. Ng, "On the cramer-rao bound under parametric constraints," *IEEE Signal Processing Letters*, vol. 5, pp. 177–179, Jul. 1998.
- [88] Z. Ben-Haim and Y. Eldar, "On the constrained cramer-rao bound with singular fisher information matrix," *IEEE Signal Processing Letters*, vol. 16, pp. 453–456, June 2009.
- [89] G. Riba, G. Vazquez, and S. Calvo, "Conditional maximum likelihood frequency estimation for offset modulation," in *IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP)*, vol. 6, May 1998, pp. 3425–3428.
- [90] R. Imad, S. Houcke, and M. Ghogho, "Blind estimation of the phase and carrier frequency offset for LDPC coded systems," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, 2010.
- [91] T. Heskes, O. Zoeter, and W. Wiegierinck, *Advances in Neural Information Processing*. MIT Press, 2004, ch. Approximate Expectation-Maximization.
- [92] N. Noels, C. Herzet, A. Dejonghe, V. Lottici, H. Steendam, M. Moeneclaey, M. Luise, and L. Vandendorpe, "Turbo-synchronization: an em approach," in *IEEE International Conference on Communications (ICC)*, 2003, pp. 2933–2937.
- [93] M. Speth, S. Fechtel, G. Fock, and M. H., "Optimum receiver design for wireless broad-band systems using OFDM - Part I," vol. 47, no. 11, pp. 1668–1677, Nov. 1999.
- [94] G. B. Giannakis, "Filterbanks for blind channel identification and equalization," vol. 4, no. 6, pp. 184–187, Jun. 1997.
- [95] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filter-bank precoders and equalizers. Part I. Unification and optimal designs and Part II. Blind channel estimation, synchronization and direct equalization," vol. 47, no. 7, pp. 1988–2022, Jul. 1999.
- [96] B. Muquet, Z. Wang, G. Giannakis, M. de Courville, and P. Duhamel, "Cyclic prefixing or zero padding for wireless multicarrier transmissions?" vol. 50, no. 12, pp. 2136–2148, Dec. 2002.
- [97] L. Hanzo, M. Muenster, B.-J. Choi, and T. Keller, *OFDM and MC-CDMA for broad-band multi-user communications*. JohnWiley & Sons, New York, 2003.
- [98] M. Morelli, "Timing and frequency synchronization for the uplink of an OFDMA system," vol. 52, no. 2, pp. 296–306, Feb. 2004.
- [99] Y. Mostofi and D. C. Cox, "Mathematical analysis of the impact of timing synchronization errors on the performance of an OFDM system," vol. 54, no. 2, pp. 226–230, Feb. 2006.

-
- [100] X. Ma, H. Kobayashi, and S. Schwartz, "Effect of frequency offset on BER of OFDM and single carrier systems," in *IEEE International Symposium on Personal Indoor and Mobile Radio Communication Proceedings*, vol. 3, 7-10 Sep. 2003, pp. 2239–2243.
- [101] K. Sathananthan and C. Tellambura, "Probability of error calculation of OFDM systems with frequency offset," vol. 49, no. 11, pp. 1884–1888, Nov. 2001.
- [102] J. J. Van de Beek, M. Sandell, and P. O. Borjesson, "ML estimation of timing and frequency offset in OFDM systems," vol. 45, no. 6, pp. 1800–1805, Jul. 1997.
- [103] M. Speth, S. Fechtel, G. Fock, and H. Meyr, "Optimum Receiver Design for OFDM-Based Broadband Transmission - Part II: A Case Study," vol. 49, no. 4, pp. 571–578, Apr. 2001.
- [104] D. Lee and K. Cheun, "Coarse Symbol Synchronization Algorithms for OFDM Systems in Multipath Channels," vol. 6, no. 10, pp. 446–448, Oct. 2002.
- [105] T. M. Schmidl and D. C. Cox, "Robust Frequency and Timing Synchronization for OFDM," vol. 45, no. 12, pp. 1613–1621, Dec. 1997.
- [106] H. Minn, V. K. Bhargava, and K. B. Letaief, "A Robust Timing and Frequency Synchronization for OFDM Systems," vol. 2, no. 4, pp. 822–839, Jul. 2003.
- [107] K. Shi and E. Serpedin, "Coarse frame and carrier synchronization of OFDM systems: A new metric and comparison," vol. 3, no. 4, pp. 1271–1284, Jul. 2004.
- [108] M. Ghogho, P. Ciblat, A. Swami, and P. Bianchi, "Training Design for Repetitive-Slot-based CFO estimation in OFDM," vol. 57, no. 12, pp. 4958–4964, Forthcoming 2009.
- [109] P. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," vol. 42, no. 10, pp. 2908–2914, Oct. 1994.
- [110] M. Ghogho and A. Swami, "Frame and frequency acquisition for OFDM," vol. 15, pp. 605–608, Dec. 2008.
- [111] M. Morelli and U. Mengali, "An improved frequency offset estimator for OFDM applications," vol. 3, no. 3, pp. 75–77, Mar. 1999.
- [112] M. Ghogho and A. Swami, "Unified framework for a class of frequency-offset estimation techniques for OFDM," in *Int. Conf. Acoustics, Speech and Signal Processing, 2004, IEEE*, May 2004.
- [113] F. Tufvesson, O. Edfors, and M. Faulkner, "Time and frequency synchronization for OFDM using PN-sequence preambles," in *Vehicular Technology Conference, 1999, IEEE*, vol. 4, Sep. 1999, pp. 2203–2207.
- [114] D. Landstrom, S. K. Wilson, J. J. Beek, P. Odling, and P. O. Borjesson, "Symbol Time Offset Estimation in Coherent OFDM Systems," vol. 50, no. 4, pp. 545–549, Apr. 2002.

-
- [115] M. Speth, F. Classen, and H. Meyr, "Frame synchronization of OFDM systems in frequency selective fading channels," in *Vehicular Technology Conference, 1997, IEEE, 1997*, pp. 1807–1811.
- [116] B. Yang, K. B. Letaief, R. S. Cheng, and Z. Cao, "Timing Recovery for OFDM Transmission," vol. 18, no. 11, pp. 2278–2291, Nov. 2000.
- [117] N. Chen, M. Tanaka, and R. Heaton, "OFDM timing synchronisation under multipath channels," in *Vehicular Technology Conference, 2003, IEEE, vol. 1, Apr. 2003*, pp. 378–382.
- [118] H. Zhou and Y.-F. Huang, "A maximum likelihood fine timing estimation for wireless OFDM systems," vol. 55, no. 1, pp. 31–41, Mar. 2009.
- [119] Y. Mostofi and D. C. Cox, "A robust timing synchronization design in OFDM systems" Cpart I: low-mobility cases," vol. 6, no. 12, pp. 4329–4339, Dec. 2007.
- [120] ———, "A robust timing synchronization design in OFDM systems" Cpart II: high-mobility cases," vol. 6, no. 12, pp. 4340–4348, Dec. 2007.
- [121] F. Daffara and O. Adami, "A novel carrier recovery technique for orthogonal multicarrier systems," *Eur. Trans. Telecommun.*, vol. 7, pp. 323–334, Jul./Aug. 1996.
- [122] F. Daffara and A. Chouly, "Maximum likelihood frequency detectors for orthogonal multicarrier systems," in *Int. Conf. on Communications, 1993, IEEE, Jun. 1993*, pp. 766–771.
- [123] M. Morelli, A. N. Andrea, and U. Mengali, "Feedback frequency synchronization for OFDM applications," vol. 5, no. 1, pp. 28–30, Jan. 2001.
- [124] H. Liu and U. Tureli, "A high-efficiency carrier estimator for OFDM communications," vol. 2, no. 4, pp. 104–106, Apr. 1998.
- [125] U. Tureli, H. Liu, and M. D. Zoltowski, "OFDM blind carrier offset estimation: ESPRIT," vol. 48, no. 9, pp. 1459–1461, Sep. 2000.
- [126] G. Santella, "A frequency and symbol synchronization system for OFDM signals: Architecture and simulation results," vol. 49, no. 1, pp. 245–275, Jan. 2000.
- [127] X. Ma, C. Tepedelenlioglu, G. B. Giannakis, and S. Barbarossa, "Non-data-aided carrier offset estimator for OFDM with null subcarriers," vol. 19, no. 12, pp. 2504–2515, Dec. 2001.
- [128] M. Ghogho, A. Swami, and G. B. Giannakis, "Optimized null-subcarrier selection for CFO estimation in OFDM over frequency selective fading channels," in *IEEE Globecom 2001*, vol. 48, Nov. 2001, pp. 202–206.
- [129] H. Bolcskei, "Blind estimation of symbol timing and carrier frequency offset in wireless OFDM systems," vol. 48, no. 6, pp. 988–999, Jun. 2001.

-
- [130] U. Tureli, D. Kivanc, and H. Liu, "Experimental and analytical studies on high-resolution OFDM carrier frequency offset estimator," vol. 50, pp. 629–643, Mar. 2001.
- [131] M. Ghogho and A. Swami, "Semi-blind frequency offset synchronization for OFDM," in *IEEE ICASSP 2002*, vol. 3, May 2002, pp. 2333–2336.
- [132] —, "Blind frequency-offset estimator for OFDM systems transmitting constant-modulus symbols," vol. 6, no. 8, pp. 343–345, Aug. 2002.
- [133] F. Yang, K. H. Li, and K. C. Teh, "A carrier frequency offset estimator with minimum output variance for OFDM systems," vol. 8, no. 11, pp. 677–679, Nov. 2004.
- [134] Y. W. Yao and G. B. Giannakis, "Blind carrier frequency offset estimation in SISO, MIMO, and multiuser OFDM systems," vol. 53, no. 1, pp. 173–183, Jan. 2005.
- [135] D. Rife and R. Boorstyn, "Single-tone parameter estimation from discrete-time observations," vol. 20, no. 5, pp. 591–598, Sep. 1974.
- [136] J. Van de Beek, P. Borjesson, M. Boucheret, D. Landstrom, J. Arenas, P. Odling, C. Ostberg, M. Wahlqvist, and S. Wilson, "A time and frequency synchronization scheme for multiuser OFDM," vol. 17, no. 11, pp. 1900–1914, Nov. 1999.
- [137] "IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems," *IEEE Std 802.16-2009 (Revision of IEEE Std 802.16-2004)*, pp. C1–2004, 29 2009.
- [138] D. Huang and K. B. Letaief, "An interference-cancellation scheme for carrier frequency offsets correction in OFDMA systems," *IEEE Transactions on Communications*, vol. 53, no. 7, pp. 1155–1165, Jul. 2005.
- [139] D. Marabissi, R. Fantacci, and S. Papini, "Robust multiuser interference cancellation for OFDM systems with frequency offset," vol. 5, no. 11, pp. 3068–3076, Nov. 2006.
- [140] S. Manohar, D. Sreedhar, V. Tikiya, and A. Chockalingam, "Cancellation of multiuser interference due to carrier frequency offsets in uplink OFDMA," vol. 6, no. 7, pp. 2560–2571, Jul. 2007.
- [141] S.-W. Hou and C. C. Ko, "Multiple-access interference suppression for interleaved OFDMA system uplink," vol. 57, no. 1, pp. 194–205, Jan. 2008.
- [142] S. Barbarossa, M. Pompili, and G. B. Giannakis, "Channel-independent synchronization of orthogonal frequency division multiple access systems," vol. 20, no. 2, pp. 474–486, Feb. 2002.
- [143] M. Morelli, C.-C. J. Kuo, and M.-O. Pun, "Synchronization techniques for orthogonal frequency division multiple access (OFDMA): a tutorial review," *Proceedings of the IEEE*, vol. 95, no. 7, pp. 1394–1427, Jul. 2007.
- [144] Z. Cao, U. Tureli, and Y. Yao, "Deterministic multiuser carrier-frequency offset estimation for interleaved OFDMA uplink," vol. 52, no. 9, pp. 1585–1594, Sep. 2004.

-
- [145] D. S. Radovic and M. M. Eric, "Performance of Subspace Based Multi-user CFO Estimation for Interleaved OFDMA Uplink," in *Proc. Int. Conf. on Computer as a Tool*, vol. 2, Belgrade, Serbia and Montenegro, Nov. 2005, pp. 1602–1605.
- [146] J. Lee, S. Lee, K.-J. Bang, S. Cha, and D. Hong, "Carrier Frequency Offset Estimation Using ESPRIT for Interleaved OFDMA Uplink Systems," vol. 56, no. 5, pp. 3227–3231, Sep. 2007.
- [147] M.-O. Pun, M. Morelli, and C.-C. Kuo, "Maximum-likelihood synchronization and channel estimation for OFDMA uplink transmissions," vol. 54, no. 4, pp. 726–736, Apr. 2006.
- [148] X. Y. Fu, Y. H. Li, and H. Minn, "A new ranging method for OFDMA systems," vol. 6, no. 2, pp. 659–669, Feb. 2007.
- [149] J.-H. Lee and S.-C. Kim, "Time and frequency synchronization for OFDMA uplink system using the SAGE algorithm," vol. 6, no. 4, pp. 1176–1181, Apr. 2007.
- [150] M.-O. Pun, M. Morelli, and C.-C. Kuo, "Iterative detection and frequency synchronization for OFDMA uplink transmissions," vol. 6, no. 2, pp. 629–639, Feb. 2007.
- [151] M. Morelli and U. Mengali, "Carrier-frequency estimation for transmissions over selective channels," vol. 48, no. 9, pp. 1580–1589, Sep. 2000.
- [152] S. Sezginer and P. Bianchi, "Asymptotically efficient reduced complexity frequency offset and channel estimators for uplink MIMO-OFDMA systems," vol. 56, no. 3, pp. 964–979, Mar. 2008.
- [153] X. N. Zeng and A. Ghayeb, "Joint cfo and channel estimation for OFDMA uplink: an application of the variable projection method," vol. 8, no. 5, pp. 2306–2311, May 2009.
- [154] Q. Z. B.R. Hamilton, X. Ma and J. Xu, "Aces: Adaptive clock estimation and synchronization using kalman filtering," in *Proc. MobiCom'08*, 2008.
- [155] L. G. J. Elson and D. Estrin, "Fine-grained network time synchronization using reference broadcasts," in *Proc. Fifth ACM SIGOPS Operating Syst. Review*, Dec 2002, pp. 147–163.
- [156] K. Romer, "Time synchronization in ad hoc networks," in *Proc. 2nd ACM Intl. Symp. Mobile Ad Hoc Networking and Comp.*, 2002, pp. 173–182.
- [157] M. Sichitiu and C. Veerarittiphan, "Simple, accurate time synchronization for wireless sensor networks," in *Proc. IEEE Wireless Comm. and Networking Conf(WCNC)*, 2003.
- [158] R. K. S. Ganeriwal and M. B. Srivastava, "Timing-sync protocol for sensor networks," in *Proc. First Intl. Conf on Embedded Networked Sensor Syst. (SenSys-03)*, 2003.

-
- [159] J. van Greunen and J. Rabaey, "Lightweight time synchronization for sensor networks," in *Proc. of WSNA'03*, 2003.
- [160] S. Graham and P. R. Kumar, "Time in general-purpose control systems: The control time protocol and an experimental evaluation," in *Proc 43rd IEEE Conf on Decision and Control*, 2004, pp. 4004–4009.
- [161] F. Cristian, "A probabilistic approach to distributed clock synchronization," *Distributed Computing*, vol. 3, pp. 146–158, 1989.
- [162] K. Arvind, "Probabilistic clock synchronization in distributed systems," *IEEE Trans. Parallel and Distributed Systems*, vol. 5, no. 5, pp. 474–487, May 1994.
- [163] J. H. D. Dolev and H. R. Strong, "On the possibility and impossibility of achieving clock synchronization," in *Proc. 16th Annual ACM Symp. on Theory of Computing*, 1984.
- [164] J. Lundelius and N. Lynch, "An upper and lower bound for clock synchronization," *Information and Control*, vol. 62, pp. 190–204, 1984.
- [165] S. Biaz and J. L. Welch, "Closed form bounds for clock synchronization under simple uncertainty assumptions," *IPL: Information Processing Letters*, vol. 80, 2001.
- [166] D. L. Mills, "Internet time synchronization: the network time protocol," *IEEE Trans. Commun.*, vol. 39, no. 10, pp. 1482–1493, Oct. 1991.
- [167] D. R. Jeske, "On maximum-likelihood estimation of clock offset," *IEEE Trans. Comm.*, vol. 53, no. 1, pp. 53–54, 2005.
- [168] A. S. D. R. Jeske, "Estimation of clock offset using bootstrap bias-correction techniques," *Technometrics*, vol. 45, no. 3, pp. 256–261, 2003.
- [169] E. S. Q.M. Chaudhari and K. Qaraqe, "On maximum likelihood estimation of clock offset and skew in networks with exponential distribution," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, April 2008.
- [170] Mirollo and S. Strogatz, "Synchronization of pulse-coupled biological oscillators," *SIAM Journal of Applied Math.*, 1990.
- [171] D. Lucarelli and I.-J. Wang, "Decentralized synchronization protocols with nearest neighbor communications," in *Proc. ACM SenSys*, 2004.
- [172] E. Mallada and A. Tang, "Synchronization of phase-coupled oscillators with arbitrary topology," in *Proc. American Control Conference (ACC'10)*, 2010.
- [173] Y.-W. Hong and A. Scaglione, "A scalable synchronization protocol for large sensor networks and its applications," *IEEE JSAC*, vol. 23, no. 5, 2005.
- [174] Q. Li and D. Rus, "Global clock synchronization in sensor networks," *IEEE Trans. Computers*, vol. 55, no. 2, pp. 214–226, 2006.

-
- [175] A. S. P. Barooah, J. Hespanha, “On the effect of asymmetric communication on distributed time synchronization,” in *Proc. 46th IEEE Conference on Decision and Control (CDC’07)*, 2007.
- [176] A. S. S. Sardellitti, S. Barbarossa, “Optimal topology control and power allocation for minimum energy consumption in consensus networks,” *IEEE Trans. on Signal Processing*, 2012.