Generalized Method of Moments Estimation for QAM Carrier Acquisition:

an Image Processing Approach

Stefano Rinauro

Dottorato di Ricerca in Ingegneria dell'Informazione e della Comunicazione

Universitá degli Studi di Roma "La Sapienza"

Dipartimento INFOCOM

A thesis submitted for the degree of

Doctor of Philosophy

Advisor: Prof. R. Cusani

He conocido lo que ignoran los Griegos: la incertidumbre

J. L. Borges

Queste poche righe per dedicare
questo lavoro ad Alessandro e Stefano che mi hanno sempre “inteso”
a Francesco e alla nostra raffinata elite squisitamente culturale
e a Sara, il mio paradiso

Acknowledgements

There is no substitute for an helpful and dedicated advisor.
I had the astonishing luck and the staggering experience of having three of them. To Profs. Cusani, Scarano and Colonnese, my deepest gratitude for everything.

Also my thanks to Prof. Giannakis, who guided me through my American experience.

And of course to my parents for their loving support.

Introduction

The *Generalized Method of Moments* (GMM) has been first formalized by Hansen in 1982 (1) and has been applied in many contexts as, for instance, economics and finance.

The GMM is an extension of the well known Method of Moments, where the estimation of a set of parameters is conducted by equating parameter dependent theoretical moments of the statistical model with their sample counterparts. The parameter's estimates are then obtained as the solution of the resulting system of equations. When either the number of equations exceeds the number of parameters to be estimated, and generalized moments are considered instead of canonicals moments (that is, the expectation of a generic function of the random variable, rather than the integer powers of the random variable itself), the method is currently referred to as the Generalized Method of Moments. Under these conditions the system is overdetermined and generally has no solution. The estimation is then conducted by minimizing a suitable weighted quadratic form. The weighting affects the performance of the estimation; when the optimal weights minimizing the estimation variance are employed, the method is referred to as the *best* GMM estimation. The optimal weights depend on the true value of the parameter. The estimated values are then obtained either by a direct solution of the minimization problem, or via an iterative approach. However, if the statistical model is nonlinear, in either case numerical algorithms have to be envisaged to solve the problem. An useful review of the GMM is found in (2; 3).

Unlike other estimation techniques, as Maximum Likelihood Estimation (MLE), GMM does not require the complete knowledge of the statistical description of the data. In fact, only specific moments derived by the underlying stochastic model are needed to perform the estimation. Besides, also in particular cases in

which the probability density function (pdf) of the data is known, MLE can be computationally unaffordable while GMM provides, under mild assumptions, a consistent estimation at a reduced complexity.

Generally, no assumptions on the efficiency of GMM estimation can be made; however, the GMM estimation coincides with MLE when the data are such that the central limit theorem applies for the selected sufficient statistic.

In the signal processing community the optimally weighted GMM estimation is currently referred to as Nonlinear Weighted Least Squares (NWLS) estimation, and it has been applied in many contexts involving, for instance, high order statistical analysis. In (4), the unweighted estimation – that is the counterpart of the unweighted GMM– is introduced, with the weight matrix replaced by the identity matrix; the asymptotically optimal formulation of this method has been introduced in (5), (6) and in (7) where the NWLS with the optimal weights matrix is employed in the case of ARMA parameter estimation with missing observations. In (8) NWLS is employed to address the problem of estimating the frequency of a complex harmonic in the presence of additive and multiplicative noise. An useful review of NWLS can be found in (9).

In this work we address the problem of location parameter estimation via a GMM approach, that is the problem of estimating a parameter that determines the location or shift of the distribution of the measurements. This problem is encountered in many applications, ranging from time-delay estimation (10) to directional statistics analysis. Examples of problems involving directional statistics can be found in several scientific disciplines as for instance in:

- *Earth Science*, for instance in the estimation of the direction of earthquake's waves or in the estimation of relative rotations of tectonic plates
- *Meteorology*, for instance concerning the estimation of wind or thunderstorms directions.
- *Medicine*, for instance about the study of the incidence of particular disease at various times of the year, or in vector cardiograms, where information about the electrical activity in a heart during a heartbeat is described in terms of a near-planar orbit in three-dimensional space.

- *Image Analysis*, for instance in machine vision related problems.

A complete review of these scenarios can be found in (11).

The main contribution of this work is to show that, when the *estimandum* acts as a shift on a suitable statistic of the underlying model, the GMM optimization problem can be modified so that the estimation is realized by means of a coarse-to-fine estimation approach using a fast, DFT based, computationally efficient procedure.

We will mathematically formalize the shift property of a parameter, by setting a condition to be satisfied by the pdf of the observations, and we will show that suitable transformations of the observed random variables can recast a non-shift problem in one exhibiting the desired property.

Besides, we show that the shift nature of the parameter can be usefully exploited to devise MLE, also when no Gaussianity assumption on the statistic can be made. In fact we show that, when the chosen statistic can be regarded as a multinomial distributed random variable, direct maximization of the likelihood function can be conducted by exactly the same estimation rule of the *unweighted* GMM. With respect to the *best* GMM estimation, a logarithmic nonlinearity is employed instead of the optimal weight matrix, thus yielding a significantly reduced computational complexity. Again a computationally efficient solution is provided via a coarse-to-fine DFT based approach.

We finally remark that, given the bind between the GMM approach and the NWLS outlined before, the solution here introduced is then feasible in all of the contexts where the NLWS estimation is employed, providing a computationally efficient solution, once a statistic affected by the parameter as a cyclic-shift is found.

Here, we analyze a GMM based gain-control-free blind carrier acquisition for QAM constellations. The carrier acquisition problem is common to many bandpass digital communication systems. In fact, even when the communication channel is ideal, *i.e.*, it solely attenuates and delays the transmitted signal, the random carrier phase rotation and frequency drift between local oscillators generates an unknown phase/frequency offset that must be recovered resorting

to trained or blind methods, depending on the application requirements. Although many standard communication systems adopt trained transmission, great bandwidth savings are achieved when the estimation is performed using blind estimators. The phase/frequency offset estimation is a well known topic in communication theory. However, existing carrier acquisition techniques require a preliminary gain adjustment stage. On the other hand, although assuming perfect gain knowledge it is possible to approach CRB in high SNR, as shown by some results presented in (12), the accuracy of gain estimation does affect both the estimation performance and the computational complexity of the estimation process. Hence there's still interest in developing gain control free estimators preserving the slope of Cramer-Rao bound. Moreover let us remark that in a carrier acquisition stage, the gain factor is typically unknown, and hence, decision directed estimation as well as gain dependent state of the art techniques result unfeasible.

Here we propose a GMM based gain-control-free estimation procedure for blind carrier acquisition for QAM constellations. The estimators are first developed from an analysis of the form of the constellations diagrams in presence of phase and frequency offset. By resorting to an image-processing driven approach, we formalized a class of estimators accounting for these regularities. More specifically, employing typical image processing tools, as for instance the tomographic projection, we devised a statistic of the observations to perform the estimation with. This class of estimators is then correctly recast in the GMM framework, providing an example of the location parameter estimation procedure introduced in this thesis. Theoretical performance analysis, assessed by numerical simulations, shows a significant performance improvement with respect to state-of-the art estimators.

This thesis is organized as follows: in Chapter 1 the Generalized Method of Moments (GMM) is briefly introduced, while in Chapter 2 the GMM estimation procedure for shift parameter is detailed. In Chapters 3-4 an image processing driven GMM-based blind phase and frequency offset acquisition for QAM constellations is introduced. Finally in Chapter 5 we report the form of the Cramér Rao Bound for phase and frequency offset gain control free estimation. Chapter 6 concludes this thesis.

Contents

Chapter 1

Generalized Method of Moments (GMM) Estimation

1.1 Introduction

In this chapter we will briefly outline the theory of Generalized Method of Moments (GMM) estimation. Unlike other estimation techniques, as Maximum Likelihood Estimation (MLE), GMM does not require the complete knowledge of the statistical description of the observations. In fact, only specific moments derived by the underlying stochastic model are needed to perform the estimation. Besides, even if the probability density function (pdf) of the data is known, MLE can be computationally unaffordable while GMM provides, under mild assumptions, a consistent estimation at a reduced complexity.

1.2 Method of Moments Estimation

Let us consider an m-dimensional parameter $\alpha = \{\alpha_1, \cdots \alpha_m\}$ that characterizes the distribution a related random variable x, and let $p_{x|\alpha}(x|\alpha)$ denote the pdf of the r.v x. The k-th moment of the r.v. x, provided it exists, is defined as follows:

$$\mu_x^{(k)} \stackrel{\text{def}}{=} \mathrm{E}\left\{x^k\right\} = \int_{-\infty}^{\infty} x^k p_{x|\alpha}(x|\alpha)\, dx \tag{1.1}$$

Note that the moments in (1.1) depend on the value of the parameter α. Let us explicitly define the relation between the k-th moment $\mu_x^{(k)}$ and the parameter α via the function $f_k(\alpha_1, \ldots, \alpha_m)$, that is $\mu_x^{(k)} = f_k(\alpha_1, \ldots, \alpha_m)$.

The moments in (1.1) can be estimated, after a number of realizations of x are observed. Let us then consider a finite number N of independent realizations of x $\{x_0, \ldots, x_{N-1}\}$; the sample estimates of $\mu_x^{(k)}$ are obtained by averaging over the observed sample as:

$$\hat{\mu}_x^{(k)} \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=0}^{N-1} x_n^k \tag{1.2}$$

Then the Method of Moments (MoM) estimation consists of equating the first m sample averaged values $\hat{\mu}_x^{(k)}$ and the first m theoretical moments in (1.1).

$$\begin{cases} \mu_x^{(1)} = & f_1(\alpha_1, \ldots, \alpha_m) = \hat{\mu}_x^{(1)} \\ \vdots & \vdots \qquad\qquad\qquad \vdots \\ \mu_x^{(m)} = & f_m(\alpha_1, \ldots, \alpha_m) = \hat{\mu}_x^{(m)} \end{cases} \tag{1.3}$$

The sample moments in (1.2), are well known to be unbiased estimates of $\mu_x^{(k)}$, provided these latter exist; moreover, by the Law of Large Numbers we have that p.lim $\hat{\mu}_x^{(k)} = \mu_x^{(k)}$. It can be proved that (13), if the inverse of the $f_k(\alpha_1, \ldots, \alpha_m)$'s, that is

$$\alpha_k = g_k\left(\mu_x^{(1)}, \cdots, \mu_x^{(m)}\right) \qquad k = 1, \cdots m$$

are continuous in $\{\mu_x^{(1)}, \cdots, \mu_x^{(m)}\}$, then the solution of (1.3) is a consistent estimator of α. The MoM hence provides, under mild conditions, a consistent estimate, also in particular cases when the pdf of the data is not known, or computationally untractable, provided that some theoretical moments are known *a priori*. In the following Section we will extend these concepts in the sense of the Generalized Method of Moments

1.3 Generalized Method of Moments Estimation

Let us consider an m-dimensional parameter α to be estimated after the observation of a finite number N of realizations of a related n-dimensional random

variable x, gathered in the vector $\mathbf{x} = [x_0, \ldots, x_{N-1}]^{\mathrm{T}}$. Let $p_{\mathrm{x}|\alpha}(\mathbf{x}|\alpha)$ denote the pdf of the observation vector $\mathbf{x}$, and $\hat{\mathbf{f}}$ a K-dimensional statistic such that $\mathrm{E}\left\{\hat{\mathbf{f}}\right\} = \mathbf{f}(\alpha)$. From the zero-mean *generalized moments*

$$\mathbf{e}(\alpha) = \hat{\mathbf{f}} - \mathbf{f}(\alpha)$$

we obtain an estimate $\hat{\alpha}$ by solving the so-called *moment conditions*:

$$\mathbf{e}(\hat{\alpha}) = \mathbf{0} \tag{1.4}$$

The MoM is obtained by this formulation when $K = m$ and by setting

$$\mathbf{e}(\alpha) \stackrel{\text{def}}{=} \left[\left(\hat{\mu}_x^{(k)} - \mu_x^{(k)}\right)\right]_{k=0}^{K-1} \tag{1.5}$$

With respect to the MoM, the Generalized Method of Moments allows to employ any statistic, and not only the canonical moments in (1.1); generally speaking, the k-th element of the vector $\mathbf{e}(\alpha)$, namely $e_k(\alpha)$, is the error between the expectation of any function of $\mathbf{x}$ and its sample average:

$$e_k(\alpha) = \mathrm{Av}\{y_k(\mathbf{x})\} - \mathrm{E}\{y_k(\mathbf{x})\} \tag{1.6}$$

wherease in the canonical MoM we would have

$$e_k(\alpha) = \mathrm{Av}\left\{x^k\right\} - \mathrm{E}\left\{x^k\right\}$$

where we compactly denoted with the operator $\mathrm{Av}\{\cdot\}$ the sample average operation, *i.e.* $\mathrm{Av}\{\mathbf{x}\} = 1/N \sum_{n=0}^{N-1} x_n$. Moreover, always with respect to the MoM, it is possible for the relations to exceed the number of paramters to be estimated.

When $K > m$ the system is overdetermined and, generally, (1.4) has no unique solution. In such a case, referred to as *Generalized Method of Moments* (GMM), an estimation of α is obtained by minimizing the following elliptic norm of $\mathbf{e}(\alpha)$:

$$\hat{\alpha}_G^{(\mathbf{W})} = \arg\min_{\alpha} \mathbf{e}(\alpha)^{\mathrm{T}} \mathbf{W}\, \mathbf{e}(\alpha) \tag{1.7}$$

being $\mathbf{W}$ a symmetric positive definite weight matrix.

It is worth pointing out that the GMM estimate $\hat{\alpha}_G^{(\mathbf{W})}$ depends on the matrix $\mathbf{W}$; it can be proved (2) that, under mild regularity conditions, the GMM estimate

$\hat{\alpha}_G$ is Consistent, that is $\text{p.lim}_{N\to\infty}\hat{\alpha}_G = \alpha$, and Asymptotically Normal (CAN) for all positive definite weight matrices $\mathbf{W}$.

Let us in fact assume that the GMM estimate $\hat{\alpha}_G^{(\mathbf{W})}$ is close enough to the true value α_0 so that we can aproximate $\mathbf{e}(\alpha)$ using a first order Taylor expansion around the true value α_0.

$$\mathbf{e}(\alpha) \simeq \mathbf{e}(\alpha_0) + \mathbf{G}_K(\alpha - \alpha_0)$$

where we denoted by $\mathbf{G}_K$ the $[K \text{ x } m]$ gradient matrix of first derivatives $\mathbf{G}_K = \partial\mathbf{e}(\alpha)/\partial\alpha^{\mathrm{T}}$. Relying on this approximation, we can re-write the cost function in (1.7) as:

$$\begin{aligned}\mathbf{e}(\alpha)^{\mathrm{T}}\mathbf{W}\,\mathbf{e}(\alpha) &\simeq (\mathbf{e}(\alpha_0) + \mathbf{G}_K(\alpha-\alpha_0))^{\mathrm{T}}\mathbf{W}(\mathbf{e}(\alpha_0) + \mathbf{G}_K(\alpha-\alpha_0)) \\ &= \mathbf{e}(\alpha_0)^{\mathrm{T}}\mathbf{W}\mathbf{e}(\alpha_0) + \mathbf{e}(\alpha_0)^{\mathrm{T}}\mathbf{W}\mathbf{G}_K(\alpha-\alpha_0) \\ &\quad + (\alpha-\alpha_0)^{\mathrm{T}}\mathbf{G}_K^{\mathrm{T}}\mathbf{W}\mathbf{e}(\alpha_0) + (\alpha-\alpha_0)^{\mathrm{T}}\mathbf{G}_K^{\mathrm{T}}\mathbf{W}\mathbf{G}_K(\alpha-\alpha_0)\end{aligned}$$

To get the minimizer of the cost function, we have to evaluate and equate to zero its first derivative:

$$\frac{\partial}{\partial\alpha}\left[\mathbf{e}(\alpha)^{\mathrm{T}}\mathbf{W}\,\mathbf{e}(\alpha)\right] \simeq \mathbf{e}(\alpha_0)^{\mathrm{T}}\mathbf{W}\mathbf{G}_K + \mathbf{G}_K^{\mathrm{T}}\mathbf{W}\mathbf{e}(\alpha_0) + 2\mathbf{G}_K^{\mathrm{T}}\mathbf{W}\mathbf{G}_K(\alpha-\alpha_0)$$

By simple algebra calculation we have:

$$\hat{\alpha}_G^{(\mathbf{W})} = \alpha_0 + (\mathbf{G}_K^{\mathrm{T}}\mathbf{W}\mathbf{G}_K)^{-1}\mathbf{G}_K^{\mathrm{T}}\mathbf{W}\,\mathbf{e}(\alpha_0) \tag{1.8}$$

Stemming from the definition of (1.6), we have that $\text{p.lim}\,\mathbf{e}(\alpha_0) = 0$, form which we can conclude that:

$$\text{p.lim}\,\hat{\alpha}_G^{(\mathbf{W})} = \alpha_0$$

which proves the consistecy of the estimator $\hat{\alpha}_G^{(\mathbf{W})}$. Moreover, assuming that the observations are such that the Central Limit Theorem applies, that is, asymptotically

$$\mathbf{e}(\alpha) \sim \mathcal{N}(\mathbf{0}, \Sigma_e)$$

we have that $\hat{\alpha}_G^{(\mathbf{W})}$ is asymptotically Gaussian with mean α_0, *i.e.* is unbiased, and variance $\text{Var}\,\hat{\alpha}_G^{(\mathbf{W})}$, with:

$$\text{Var}\,\hat{\alpha}_G^{(\mathbf{W})} = (\mathbf{G}_K^{\mathrm{T}}\mathbf{W}\mathbf{G}_K)^{-1}\mathbf{G}_K^{\mathrm{T}}\mathbf{W}\Sigma_e\mathbf{W}\mathbf{G}_K(\mathbf{G}_K^{\mathrm{T}}\mathbf{W}\mathbf{G}_K)^{-1} \tag{1.9}$$

A summary of the asymptotic properties of the GMM estimator is found in (2).

When no weighting is comprised, *i.e.* the matrix $\mathbf{W}$ is the identity matrix ($\mathbf{W} = \mathbf{I}$), the estimator (1.7) is known as the *unweighted* GMM estimator.

The weight matrix $\mathbf{W}$ can be selected so as to minimize the estimator's variance and the resulting optimal matrix $\mathbf{W}^{(o)}$ is proved to exhibit the following form (2):

$$\mathbf{W}^{(o)} = \mathbf{W}_0(\alpha_0) = \frac{1}{N}\left[\mathrm{E}\left\{\mathbf{e}(\alpha_0)\,\mathbf{e}(\alpha_0)^{\mathrm{T}}\right\}\right]^{-1} \tag{1.10}$$

The optimal minimum variance GMM estimate is obtained by minimizing the objective function as in (1.7) with $\mathbf{W} = \mathbf{W}^{(o)}$, and it is also known as the *best* GMM estimate.

When the optimal weight matrix is employed, the asymptotic variance in (1.9) simplifies to:

$$\mathrm{Var}\,\hat{\alpha}_G^{(\mathbf{W}^{(o)})} = \left(\mathbf{G}_K^{\mathrm{T}}\Sigma_e^{-1}\mathbf{G}_K\right)^{-1} \tag{1.11}$$

If $\mathbf{e}(\alpha)$ is nonlinear, generally no closed form solution is available, and numerical algorithms have to be employed. Besides, it is worth noting that the optimal matrix depends itself on the true value α_0. The problem in (1.7) is solved either directly for α, or via a coarse-to fine approach. In this latter case a suboptimal estimate $\hat{\alpha}_G^{(\mathbf{I})}$ is first found by setting $\mathbf{W} = \mathbf{I}$.

$$\hat{\alpha}_G^{(\mathbf{I})} = \arg\min_{\alpha} \mathbf{e}(\alpha)^{\mathrm{T}}\,\mathbf{e}(\alpha)$$

According to (1.10), this intermediate estimate is then used to evaluate the optimal weight matrix $\mathbf{W}^{(o)}$ so to obtain the optimal estimate:

$$\hat{\alpha}_G^{(\mathbf{W}^{(o)})} = \arg\min_{\alpha} \mathbf{e}(\alpha)^{\mathrm{T}}\,\mathbf{W}_0\left(\hat{\alpha}_G^{(\mathbf{I})}\right)\mathbf{e}(\alpha)$$

Both the procedures require the numerical solution of minimization problems.

In the following Chapter we will show that when the *estimandum* acts as a shift on a suitable statistic of the underlying model, the GMM estimation can be realized by a coarse-to-fine computationally efficient procedure.

Chapter 2

GMM Estimation of Location Parameters

2.1 Introduction

In this Chapter, we will show how suitable design of the generalized moment $\mathbf{e}(\alpha)$ allows to derive a computationally efficient, DFT based, GMM estimation procedure in the noticeable case of α being a location parameter. We will briefly review the concept of location parameter of a pdf family; in particular, we will consider α being a location parameter either directly acting on the observed random variable x or after a suitable transformation $\mathcal{Z}(x)$. Then, we will show how a proper selection of $\mathbf{f}(\alpha)$ both directly and through its unbiased estimate $\hat{\mathbf{f}}$, allows to develop a computationally efficient, DFT based, GMM estimation procedure.

Moreover we analyze the relation between GMM and Maximum Likelihood Estimation, investigating the case of Normally distributed statistic, for which the best GMM estimation is indeed a ML estimation, and the case of Multinomially distributed statistic, for which the ML estimate is obtained by the same estimation rule of the unweighted GMM estimator.

2.2 Location Parameters

To introduce the definition of location parameter, let us consider the aforementioned n-dimensional random variable x depending on an unknown parameter (*estimandum*) α.[1] Let us also consider a transformation $\mathcal{Z} : \mathbb{K}^n \to \mathbb{R}^l$ and the corresponding transformed random variable $z = \mathcal{Z}(x)$. Then, α is said to be a location parameter for the pdf family $p_{\mathrm{z}|\alpha}\left(z^{(1)}, \ldots, z^{(l-1)}, z^{(l)}|\alpha\right)$ when this latter depends on α only through the difference $z^{(l)} - \alpha$:[2]

$$p_{\mathrm{z}|\alpha}\left(z^{(1)}, \ldots, z^{(l-1)}, z^{(l)}|\alpha\right) = p_{\mathrm{z}|\alpha}\left(z^{(1)}, \ldots, z^{(l-1)}, z^{(l)} - \alpha|0\right) \tag{2.1}$$

As an example concerning the transformation $\mathcal{Z}(x)$, a noticeable case undoubtedly regards the scale parameter estimation problem, occuring when

$$\begin{aligned} &p_{\mathrm{x}|\alpha}\left(x^{(1)}, \ldots, x^{(l-1)}, x^{(l)}|\alpha\right) \\ &\qquad = p_{\mathrm{x}|\alpha}\left(x^{(1)}, \ldots, x^{(l-1)}, \frac{x^{(l)}}{\alpha}\middle| 1\right) \end{aligned}$$

The transformation $\mathcal{Z}(x)$ realizing the exponential warping of the l-th axis, *i.e.*

$$\mathcal{Z}(x) = \begin{cases} z^{(i)} = x^{(i)} & \text{for } i = 0, \ldots, l-1 \\ \\ z^{(l)} = \log x^{(l)} \end{cases}$$

turns the scale parameter in a location one and it is very simple matter to verify that (2.1) now holds having care to substitute α with $\log \alpha$.

After these settings, let us consider the following nonlinear moment:

$$\begin{aligned} f(\xi - \alpha) = &\frac{1}{\Delta} \int_{\xi - \Delta/2}^{\xi + \Delta/2} dz^{(l)} \\ &\cdot \left[\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} h\left(z^{(1)}, \ldots, z^{(l-1)}\right) \cdot p_{\mathrm{z}|\alpha}(z|\alpha)\ dz^{(1)} \cdots dz^{(l-1)} \right] \end{aligned} \tag{2.2}$$

[1] From now on we assume for the sake of simplicity, and without loss of generality, that α is a one-dimensional parameter, *i.e.* $m = 1$; extension of the analysis to the multidimensional case is straightforward.

[2] If (2.1) holds for $\mathcal{Z}(x) = x$ then α is a location parameter for the pdf family $p_{\mathrm{x}|\alpha}\left(x^{(1)}, \ldots, x^{(n-1)}, x^{(n)}|\alpha\right)$.

being $h\left(z^{(1)},\ldots,z^{(l-1)}\right)$ an $(l-1)$-dimensional function that can be specified according to different design criteria. As a consequence of (2.1), the nonlinear moment $f(\xi-\alpha)$ depends only on $\xi-\alpha$, *i.e.* since α is a location parameter for the pdf family $p_{\mathrm{z}|\alpha}\left(z^{(1)},\ldots,z^{(l-1)},z^{(l)}|\alpha\right)$, it plays the role of a shift parameter on the nonlinear moment $f(\xi-\alpha)$.

To gain more insight on the nonlinear moment $f(\xi-\alpha)$ defined in (2.2), we observe that, when $h\left(z^{(1)},\ldots,z^{(l-1)}\right)=1$, it reduces to the normalized area of the marginal pdf $p_{\mathrm{z}^{(l)}|\alpha}\left(z^{(l)}|\alpha\right)$ over an interval of width Δ centered in ξ:

$$f(\xi-\alpha)=\frac{1}{\Delta}\int_{\xi-\Delta/2}^{\xi+\Delta/2}dz^{(l)}\cdot\left[\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}p_{\mathrm{z}|\alpha}\left(z|\alpha\right)\,dz^{(1)}\cdots dz^{(l-1)}\right]$$

On the other hand, when $h\left(z^{(1)},\ldots,z^{(l-1)}\right)$ is not constant, a weighting of the variables $(z^{(1)},\ldots,z^{(l-1)})$ is performed in carrying out the $(l-1)$-dimensional integral saturation in (2.2).

The sample estimation of $f(\xi-\alpha)$ on a set of K points $\xi_k,\ k=0,\ldots,K-1$ can be conducted by means of a weighted histogram where each occurrence of the variable $z^{(l)}$ is attributed a weight $h\left(z^{(1)},\ldots,z^{(l-1)}\right)$ depending on the co-occurrences of the variables $z^{(1)},\cdots z^{(l-1)}$. Specifically, when a sample of N statistically independent realizations z_n, for $n=1,\ldots,N$ is available, we form the following weighted histogram trough the following accumulation on intervals large Δ_k:

$$\hat{f}_k=\frac{1}{N\,\Delta_k}\sum_{n=1}^{N}h\left(z_n^{(1)},\ldots,z_n^{(l-1)}\right)\mathrm{rect}\left(\frac{z_n^{(l)}-\xi_k}{\Delta_k}\right)\tag{2.3}$$

The evaluation of the variances-covariances of the sample estimates $\hat{f}_k$ defined in (2.3) are reported in App. A.

2.3 GMM Estimation of Location Parameters

Let us assume $f(\xi)$ band-limited so that the rate $1/\Delta$ assures Nyquist sampling;[1] the uniform sampling of $f(\xi-\alpha)$ in K equispaced points $\xi_k=(2k+1)\Delta/2$, for

[1] We will assume band-limitation to be verified in all the subsequent analytical developments.

$k = 0, \ldots, K-1$, is gathered in the vector $\mathbf{f}(\alpha)$:

$$\mathbf{f}(\alpha) = \Big[f(\xi_k - \alpha) \Big]_{k=0}^{K-1}$$

Taking $\Delta_k = \Delta$ in (2.3), the corresponding unbiased estimates $\hat{f}_k$ are gathered in the vector $\hat{\mathbf{f}}$:

$$\hat{\mathbf{f}} = \Big[\hat{f}_k \Big]_{k=0}^{K-1}$$

and the following generalized moment is formed:

$$\mathbf{e}(\alpha) = \hat{\mathbf{f}} - \mathbf{f}(\alpha) \tag{2.4}$$

Stemming from the definition of $\hat{\mathbf{f}}$ and $\mathbf{f}(\alpha)$, the generalized moment defined in (2.4) satisfies the moment conditions expressed in (1.4). Hence, considering the elliptic norm:

$$\mathcal{Q}(\alpha; \mathbf{W}) = \Big[\hat{\mathbf{f}} - \mathbf{f}(\alpha) \Big]^{\mathrm{T}} \mathbf{W} \Big[\hat{\mathbf{f}} - \mathbf{f}(\alpha) \Big]$$

the GMM estimate of α is obtained as follows:

$$\hat{\alpha}_G^{(\mathbf{W})} = \arg\min_a \mathcal{Q}(a; \mathbf{W}) \tag{2.5}$$

We observe that the nonlinear moments $\mathbf{f}(\alpha)$ and $\mathbf{f}(0)$ are obtained by sampling $f(\xi - \alpha)$ and $f(\xi)$, which are equal unless the shift α, and hence the elements of $\mathbf{f}(\alpha)$ can be obtained from the elements $\mathbf{f}(0)$ by suitable interpolation. Using this property, in the following we show how GMM location parameter estimation can be obtained by a computationally efficient approach based on the DFT.

2.3.1 Unweighted GMM Estimation

Let us first consider the unweighted GMM cost function obtained for $\mathbf{W} = \mathbf{I}$:

$$\begin{aligned} \mathcal{Q}(\alpha; \mathbf{I}) &= (\hat{\mathbf{f}} - \mathbf{f}(\alpha))^{\mathrm{T}} (\hat{\mathbf{f}} - \mathbf{f}(\alpha)) \\ &= \hat{\mathbf{f}}^{\mathrm{T}} \cdot \hat{\mathbf{f}} \;\; + \mathbf{f}(\alpha)^{\mathrm{T}} \cdot \mathbf{f}(\alpha) - 2\, \hat{\mathbf{f}}^{\mathrm{T}} \cdot \mathbf{f}(\alpha) \end{aligned} \tag{2.6}$$

In order to exploit the cyclic convolution property of the DFT for the evaluation of the cost function (2.6), we observe that, since $f(\xi-\alpha)$ is band-limited, the moment $f(\xi-\alpha)$ is obtained by interpolation of $f(\xi)$ as:

$$f(\xi-\alpha)=\sum_{n=-\infty}^{\infty} f(\xi_n)\,\frac{\sin(K(\xi-\alpha-\xi_n)/2)}{K(\xi-\alpha-\xi_n)/2} \tag{2.7}$$

and the sequence of samples $f(\xi_k-\alpha)$ is obtained by low-pass filtering the sequence $f(\xi_k)$. Besides, to implement the filtering using the DFT, without loss of generality we may assume that the nonlinear moment $f(\xi)$ is periodic modulo 2π,[1] so as to have $\Delta=2\pi/K$. With this position, the sequence of samples $f(\xi_k-\alpha)$ is obtained from the sequence $f(\xi_k)$ as a cyclic convolution between the sequences of samples[2]:

$$f(\xi_k-\alpha)=\sum_{n=0}^{K-1} f(\xi_n)\,\frac{\sin(K(\xi_k-\xi_n-\alpha)/2)}{K\sin((\xi_k-\xi_n-\alpha)/2)} \tag{2.8}$$

Using a compact matrix form, we have:

$$\mathbf{f}(\alpha)=\mathbf{S}(\alpha)\cdot\mathbf{f}(\mathbf{0}) \tag{2.9}$$

the matrix $\mathbf{S}(\xi)$ being the *periodic sinc* interpolation matrix, with entries:

$$||\mathbf{S}(\alpha)||_{k,n}=\left.\frac{\sin(K\omega/2)}{K\sin(\omega/2)}\right|_{\omega=(k-n)\Delta-\alpha}$$

Notice that the matrix $\mathbf{S}(\alpha)$ is orthogonal, *i.e.* $\mathbf{S}(\alpha)^{\mathrm{T}}\mathbf{S}(\alpha)=\mathbf{I}$, and hence the squared ℓ_2 norm of the vector $\mathbf{f}(\alpha)$ does not depend on α:

$$\mathbf{f}(\alpha)^{\mathrm{T}}\cdot\mathbf{f}(\alpha)=\mathbf{f}(\mathbf{0})^{\mathrm{T}}\cdot\mathbf{S}(\alpha)^{\mathrm{T}}\mathbf{S}(\alpha)\cdot\mathbf{f}(\mathbf{0})=\mathbf{f}(\mathbf{0})^{\mathrm{T}}\cdot\mathbf{f}(\mathbf{0})$$

[1] This is always possible when $f(\xi)$ is different from zero on a finite interval, while it has to be considered in an approximate fashion when a sufficiently accurate version is obtained by windowing on the ξ axis.

[2] In fact we have

$$\sum_{k=-\infty}^{+\infty}\frac{\sin(K(\xi-k2\pi)/2)}{K(\xi-k2\pi)/2}=\frac{\sin(K\xi/2)}{K\sin(\xi/2)}$$

.

so that the minimal value of $\mathcal{Q}(\alpha;\mathbf{I})$ is found by maximizing the scalar product $\hat{\mathbf{f}}^{\mathrm{T}}\cdot\mathbf{f}(\alpha)$ w.r.t. α:

$$\hat{\alpha}_G^{(\mathbf{I})} = \arg\min_a \mathcal{Q}(a;\mathbf{I}) = \arg\max_a \ \hat{\mathbf{f}}^{\mathrm{T}}\cdot\mathbf{f}(a) \tag{2.10}$$

Using (2.9), the estimation rule is finally expressed as follows:

$$\hat{\alpha}_G^{(\mathbf{I})} = \arg\max_a \ \hat{\mathbf{f}}^{\mathrm{T}}\cdot\mathbf{S}(a)\cdot\mathbf{f}(\mathbf{0}) \tag{2.11}$$

Given the shift nature of the parameter, a fast, computationally efficient, DFT based procedure can be employed to obtain $\hat{\alpha}_G^{(\mathbf{I})}$. Let us discretize the range $[0,2\pi)$ of the cyclic parameter in K intervals of width $2\pi/K$. Then, since $\Delta = 2\pi/K$, it results:

$$\mathbf{S}(2\pi k/K) = \begin{cases} \mathbf{I} & k=0 \\ \mathbf{D}^k & k=1,\cdots K-1 \end{cases}$$

being $\mathbf{D}=(\mathbf{D}^{\mathrm{T}})^{-1}$ the following orthogonal unit cyclic shift matrix:

$$\mathbf{D} \stackrel{\mathrm{def}}{=} \begin{bmatrix} 0 & \cdots & \cdots & 0 & 1 \\ 1 & 0 & \ddots & \ddots & 0 \\ 0 & 1 & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}$$

The index k_c corresponding to the coarse estimate with resolution $2\pi/K$, namely $\hat{\alpha}_G^{(\mathbf{I})} = 2\pi k_c/K$, is obtained from (2.11) rewritten as follows:

$$k_c = \arg\min_k \mathcal{Q}\left(\frac{2\pi k}{K};\mathbf{I}\right) = \arg\max_k \ \hat{\mathbf{f}}^{\mathrm{T}}\cdot\mathbf{D}^k\cdot\mathbf{f}(\mathbf{0}) \tag{2.12}$$

Due to the cyclic shift property of the operator $\mathbf{D}$, in the rightmost hand side of (2.12) we recognize the cyclic cross-correlation between the sequences collected in the vectors $\hat{\mathbf{f}}$ and $\mathbf{f}(\mathbf{0})$; thus, the maximization over k can be conducted by properly using the DFT of both the sequences since the index k_c is nothing else that the *lag* that locates the cross-correlation maximum. The overall computational complexity is thus significantly reduced by choosing the value of K according to specific FFT algorithms.

2.3.2 Minimum Variance GMM Estimation

Let us now consider the *best* GMM estimator, *i.e.* the GMM estimator that minimizes $\mathrm{Var}\left\{\hat{\alpha}_G^{(\mathbf{W})}\right\}$ with respect to the weight matrix $\mathbf{W}$; the *best* GMM estimator weight matrix

$$\mathbf{W}^{(o)} = \arg\min_{\mathbf{W}} \mathrm{Var}\left\{\hat{\alpha}_G^{(\mathbf{W})}\right\}$$

is known to be found for $\mathbf{W}_0(\alpha) = \mathbf{\Omega}(\alpha)^{-1}/N$, being

$$\mathbf{\Omega}(\alpha) \stackrel{\mathrm{def}}{=} \mathrm{E}\left\{(\hat{\mathbf{f}} - \mathbf{f}(\alpha))(\hat{\mathbf{f}} - \mathbf{f}(\alpha))^{\mathrm{T}}\right\}$$

the covariance matrix of the measurements (2).

Since the evaluation of the inverse measurements covariance matrix $\mathbf{\Omega}(\alpha)^{-1}$ requires the knowledge of the parameter α, a coarse-to-fine approach can be envisaged for the minimization of $\mathcal{Q}(\alpha; \mathbf{W}_0(\alpha))$. Namely, once a coarse GMM estimate $\hat{\alpha}_G^{(c)}$ is found,[1] the optimal GMM cost function $\mathcal{Q}(\alpha; \mathbf{W}_0(\alpha))$ can be approximated as follows:

$$\begin{aligned}\mathcal{Q}(\alpha; \mathbf{W}_0(\alpha)) &\approx \mathcal{Q}(\alpha; \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right)) \\ &= \left[\hat{\mathbf{f}} - \mathbf{f}(\alpha)\right]^{\mathrm{T}} \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right)\left[\hat{\mathbf{f}} - \mathbf{f}(\alpha)\right]\end{aligned}$$

so as to obtain a fine estimate as follows:

$$\hat{\alpha}_G^{(\mathbf{W}^{(o)})} = \arg\min_a \mathcal{Q}\left(a; \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right)\right)$$

Since the term $\hat{\mathbf{f}}^{\mathrm{T}} \cdot \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right) \cdot \hat{\mathbf{f}}$ does not depend on α, we can also consider the following estimation rule:

$$\hat{\alpha}_G^{(\mathbf{W}^{(o)})} = \arg\max_a \mathcal{J}\left(a; \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right)\right) \tag{2.13}$$

with

$$\begin{aligned}\mathcal{J}\left(a; \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right)\right) &= \hat{\mathbf{f}}^{\mathrm{T}} \cdot \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right) \cdot \mathbf{f}(a) \\ &\quad - \frac{1}{2} \cdot \mathbf{f}(a)^{\mathrm{T}} \cdot \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right) \cdot \mathbf{f}(a)\end{aligned} \tag{2.14}$$

[1] As previously described, a coarse GMM estimate of α can be found by minimizing (2.6), *i.e.* from (2.12) and $\hat{\alpha}_G^{(c)} = \hat{\alpha}_G^{(\mathbf{I})} = 2\pi k_c/K$.

The search for the maximum can be conducted by means of a suitable interpolation after a few values of $\mathcal{J}\left(a; \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right)\right)$ have been evaluated around the coarse estimate $\hat{\alpha}_G^{(c)}$. For instance, following the approach (15), once a rough estimate $\hat{\alpha}_G^{(c)}$ Has been obtained, a parabolic approximation for the GMM cost function $\mathcal{J}\left(a; \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right)\right)$ around its maximum can be employed to calculate a fine estimate $\hat{\alpha}_G^{(f)}$, see (4.19) . The estimation form (4.19) is analytically tractable, and it allows to evaluate the asymptotical performance of the optimal GMM based estimation. Moreover, the computational cost is not significantly increased, since only few values of the objective function need to be evaluated, and a fast procedure for computing $\mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right)$ can be devised, as summarized in 2.3.3.

$$\hat{\alpha}_G^{(f)} = \hat{\alpha}_G^{(c)} - \frac{\Delta}{2} \cdot \frac{\mathcal{J}\left(\hat{\alpha}_G^{(c)}+\Delta; \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right)\right) - \mathcal{J}\left(\hat{\alpha}_G^{(c)}-\Delta; \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right)\right)}{\mathcal{J}\left(\hat{\alpha}_G^{(c)}+\Delta; \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right)\right) - 2\mathcal{J}\left(\hat{\alpha}_G^{(c)}; \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right)\right) + \mathcal{J}\left(\hat{\alpha}_G^{(c)}-\Delta; \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right)\right)} \tag{2.15}$$

2.3.3 On the Inversion of the Variance-Covariance Matrix

If the coarse estimation $\hat{\alpha}_G^{(c)}$ is chosen to be as the unweighted GMM estimation, that is the solution of (2.6), the inversion of the variance-covariance matrix $\mathbf{\Omega}\left(\hat{\alpha}_G^{(c)}\right)$ can be conducted with an extermely reduced computational complexity.

Let us infact suppose to have evaluate a coarse estimation $\hat{\alpha}_G^{(c)} = \hat{\alpha}_G^{(\mathbf{I})}$; from (2.12) we have that $\hat{\alpha}_G^{(c)} = 2\pi k_c/K$.

For $\hat{\alpha}_G^{(c)} = 2\pi k_c/K$, and since α is a cyclic location parameter, the covariance matrix factorizes as follows:

$$\mathbf{\Omega}\left(\hat{\alpha}_G^{(c)}\right) = \mathbf{D}^{k_c}\mathbf{\Omega}(0)(\mathbf{D}^{\mathrm{T}})^{k_c}$$

Stemming from the orthogonality of the matrix $\mathbf{D}$, the inverse covariance matrix $\mathbf{\Omega}\left(\hat{\alpha}_G^{(c)}\right)^{-1}$ is simply evaluated as:

$$\mathbf{\Omega}(\hat{\alpha}_G^{(c)})^{-1} = \mathbf{D}^{k_c}\mathbf{\Omega}(0)^{-1}(\mathbf{D}^{\mathrm{T}})^{k_c}$$

This relation can be exploited to pre-calculate offline and store the inverse matrix $\mathbf{\Omega}(0)^{-1}$, and to obtain online the matrix $\mathbf{\Omega}\left(\hat{\alpha}_G^{(c)}\right)^{-1}$ by columns and rows cyclic shifting.

As far as the computational load needed to compute the GMM weighted cost function in the three points $\alpha = \hat{\alpha}_G^{(c)}$ and $\alpha = \hat{\alpha}_G^{(c)} \pm \Delta = \hat{\alpha}_G^{(c)} \pm 2\pi/K$ is concerned, note that (2.14) can be written as follows:

$$\begin{aligned} &\left[\hat{\mathbf{f}} - \frac{1}{2}\mathbf{f}\left(\hat{\alpha}_G^{(c)} + q \cdot 2\pi/K\right)\right]^{\mathrm{T}} \mathbf{W}_0 \left(\hat{\alpha}_G^{(c)}\right) \mathbf{f}\left(\hat{\alpha}_G^{(c)} + q \cdot 2\pi/K\right) \\ &\qquad = \left[\mathbf{D}^{-k_c}\hat{\mathbf{f}} - \frac{1}{2}\mathbf{D}^q \mathbf{f}\,(0)\right]^{\mathrm{T}} \boldsymbol{\Omega}(0)^{-1}\, \mathbf{D}^q\, \mathbf{f}\,(0) \end{aligned} \tag{2.16}$$

where $q = 0$ indicates $\alpha = \hat{\alpha}_G^{(c)}$, and $q = \pm 1$ respectively indicates $\alpha = \hat{\alpha}_G^{(c)} \pm 2\pi/K$. Thus, from (2.16) we see that it is sufficient to pre-calculate and store only the three vectors $\boldsymbol{\Omega}(0)^{-1}\, \mathbf{D}^q\, \mathbf{f}\,(0)$, and the three scalars $\frac{1}{2}\mathbf{f}\,(0)^{\mathrm{T}}\ \mathbf{D}^{-q}\, \boldsymbol{\Omega}(0)^{-1}\, \mathbf{D}^q\, \mathbf{f}\,(0)$, while the product $\mathbf{D}^{-k_c}\hat{\mathbf{f}}$ is realized by a simple cyclic rotation of the elements of $\hat{\mathbf{f}}$. Therefore, the total number of arithmetic operations is $3K$ muls/adds.

2.4 GMM and Maximum Likelihood Estimation

In this Section we discuss the relation between GMM estimation and MLE. After having recalled that under the assumption of Guassianity of the employed statistic the optimally weighted GMM estimator coincides with the Maximum Likelihood (ML) estimator, we will show that under the assumption of multinomially distributed statistic, if the *estimandum* is a shift parameter the MLE is still obtained by the same estimation rule as the unweighted GMM estimation.

Let us consider, from now on, that $\mathbf{e}\,(\alpha)$ is a sufficient statistic for the estimation problem under examination.

2.4.1 Normally Distributed Statistic

Let us assume that the data $\mathbf{x}$ are such that the central limit theorem applies to the generalized moments (2.4):

$$\mathbf{e}\,(\alpha) \sim \mathcal{N}\,(\mathbf{0}, \Sigma_{\mathbf{e}}) \tag{2.17}$$

being $\Sigma_{\mathbf{e}}$ the variances/covaraiances matrix of $\mathbf{e}\,(\alpha)$.

Within this assumption, the ML estimate of the parameter is given by:

$$\begin{aligned} \hat{\alpha}_{ML} &= \arg\min_{\alpha} \mathcal{Q}_{ML}(\alpha) \\ \mathcal{Q}_{ML}(\alpha) &= \mathbf{e}(\alpha)^{\mathrm{T}} \Sigma_{\mathbf{e}}^{-1} \mathbf{e}(\alpha) \end{aligned} \tag{2.18}$$

It can be easily recognized that the minimization (2.18) is exactly the same of (1.7) where $\Sigma_{\mathbf{e}}^{-1}$ is employed as weight matrix in lieu of $\mathbf{W}$. Recalling that the minimum variance GMM estimation employees the optimal weight matrix

$$\mathbf{W}^{(o)} = \frac{1}{N}\left[\mathrm{E}\left\{\mathbf{e}(\alpha)\,\mathbf{e}(\alpha)^{\mathrm{T}}\right\}\right]^{-1} = \Sigma_{\mathbf{e}}^{-1}$$

it stems out that the maximization of the log-likelihood function is obtained by minimizing exactly the same objective function in (2.14). Thus, within the assumption of Gaussianity for the statistic at hand, it remains proved that the best GMM estimator and the ML estimator are the same.

2.4.2 Multinomially Distributed Statistic

Here, we show how the MLE can be conducted at a reduced computational complexity when the underlying distribution is multinomial.

Let us consider the nonlinear moment (2.2) with $h\left(z^{(1)},\ldots,z^{(l-1)}\right)=1$. and a complete class of mutually exclusive events $\mathcal{E}_k = \{|z^{(l)} - \xi_k| < \Delta/2\}$ for $k = 0,\ldots,K-1$. Hence, within a proportionality factor equal to Δ, the nonlinear moment $f(\xi_k - \alpha)$ now takes the meaning of probability of the event $\mathcal{E}_k$.

Let ν_k the number of occurrences of $\mathcal{E}_k$ in N statistically independent trials so that the probability distribution of the K random variables $\nu_0,\ldots,\nu_{K-1}$ is multinomial:

$$P(\nu_0,\ldots,\nu_{K-1}) = \frac{N!}{\prod_{k=0}^{K-1}\nu_k!} \cdot \prod_{k=0}^{K-1}\left[\Delta \cdot f(\xi_k - \alpha)\right]^{\nu_k}$$

These statistics actually represent an histogram; therefore, the herein presented multinomial model of the observations applies to any framework in which the measurements are collected in the form of histogram. For instance, this approach is quite common in communications applications, when physical measurements are collected for calibration or characterization purposes (see for instance

(16)-(18)), as well as in image processing applications, where the evaluation of the image histogram allows to compactly characterize at the first order the otherwise high-dimensional image domain, for restoration, classification, or watermarking purposes (19), (20).

Stemming from these observations, the log-likelihood of $\hat{f}_k = \nu_k / N\Delta$ can be written as:

$$\begin{aligned} l(\hat{\mathbf{f}}; \mathbf{f}(\alpha)) &= A + N\Delta \sum_{k=0}^{K-1} \hat{f}_k \ln\left(f(\xi_k - \alpha)\right) \\ A &= \ln N! - \ln \prod_{k=0}^{K-1} (N\Delta \cdot \hat{f}_k)! + N\Delta \sum_{k=0}^{K-1} \hat{f}_k \ln \Delta \end{aligned} \tag{2.19}$$

MLE is obtainend by maximizing (2.19) with respect to α:

$$\hat{\alpha}_{\mathrm{ML}} = \arg\max_{\alpha} \; l(\hat{\mathbf{f}}; \mathbf{f}(\alpha)). \tag{2.20}$$

After having compactly denoted $\tilde{\mathbf{f}}(\alpha) \stackrel{\mathrm{def}}{=} \ln \mathbf{f}(\alpha)$, and neglecting the term A in (2.19) and all the other terms in (2.20) that do not explicitly depend on α, we end up with the following ML estimation rule:

$$\hat{\alpha}_{\mathrm{ML}} = \arg\max_{\alpha} \; \hat{\mathbf{f}}^{\mathrm{T}} \cdot \tilde{\mathbf{f}}(\alpha) \tag{2.21}$$

Then, since α is a location parameter for the statistic $f(\xi_k - \alpha)$, we can rewrite the inner product in (2.21) as the cross-correlation function between the vector collecting the values of the statistic and its expected value transformed by a suitable logarithmic nonlinearity:

$$\begin{aligned} \hat{\alpha}_{\mathrm{ML}} &= \arg\max_{\alpha} \mathcal{C}(\alpha) \\ \mathcal{C}(\alpha) &= \sum_{k=0}^{K-1} \hat{f}_k \ln f(\xi_k - \alpha) \end{aligned} \tag{2.22}$$

Recalling the previous discussion concerning with the computational issues in the maximization (2.12), we conclude that the structure of the ML maximization (2.22) allows to employ a fast FFT based grid search attained at a reduced computational complexity.

Moreover, it is very interesting to notice that the estimation rule (2.22) is very close to the unweighted GMM one in (2.10), from which it differs only by

employment of the nonlinearly transformed $\tilde{\mathbf{f}}(\mathbf{0})$ instead of $\mathbf{f}(\mathbf{0})$. The difference with respect to the best GMM estimator is that now weighting is implicitly performed by the logarithmic nonlinearity, resulting in a significant reduction of the overall computational complexity.

Again $\hat{\alpha}_{\mathrm{ML}}$ is obtained in a two-stage, coarse-to-fine, estimation procedure; the coarse estimate provided by the discrete cross-correlation is definitely limited by the resolution given by the value of K; a fine estimate is attained by means of a suitable interpolation of the objective function around its maximum.

Let us finally remark that the cost function defined according to the ML criterion as in (2.20) is generally non-convex; the pursuit of its global maximum is then a nontrivial task. However the DFT approach, direct consequence of the shift nature of the *estimandum* with respect to the statistic employed, allows an efficient detection of the global maximum of the likelihood, within a resolution given by the value of K.

2.5 Performance Analysis

Here we outline the performance analysis of both the GMM based and the reduced complexity ML estimators introduced in the previous Sections.

The asymptotic estimation variance of the *best* GMM estimator is proved to be given by (1; 2):

$$N \cdot \mathrm{Var}\left\{\hat{\alpha}_G^{(\mathbf{W}^{(o)})}\right\} = \nabla_\alpha \mathbf{f}(\alpha)^{\mathrm{T}} \cdot \mathbf{W}^{(o)} \cdot \nabla_\alpha \mathbf{f}(\alpha) \tag{2.23}$$

where the gradient vector $\nabla_\alpha \mathbf{f}(\alpha)$ collects the first derivatives of nonlinear moment in (2.2) with respect to the parameter.

$$\nabla_\alpha \mathbf{f}(\alpha) \stackrel{\mathrm{def}}{=} \left[\frac{\partial f(\xi_0 - \alpha)}{\partial \alpha}, \ldots, \frac{\partial f(\xi_{K-1} - \alpha)}{\partial \alpha}\right]^{\mathrm{T}}$$

Since the analytical expression of these derivatives may be quite involved, we conduct the analytical performance evaluation resorting to the parabolic interpolation formulas as in (15), where the objective function around its maximum is approximated with the second order Taylor expansion. Following the guidelines of (15; 41), the asymptotic variance of the fine estimate $\hat{\alpha}_G^{(f)}$ in (2.15)

can be expressed as a linear combination of the variances and covariances of $\mathcal{J}\left(\xi; \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right)\right)$ for $\xi = \hat{\alpha}_G^{(c)} + m\Delta,\ m = 0, \pm 1$.

Let us set:

$$\begin{aligned} x &\stackrel{\text{def}}{=} \mathcal{J}\left(\hat{\alpha}_G^{(c)} + \Delta; \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right)\right), \quad & X &\stackrel{\text{def}}{=} \mathrm{E}\{x\}, \\ y &\stackrel{\text{def}}{=} \mathcal{J}\left(\hat{\alpha}_G^{(c)} - \Delta; \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right)\right), \quad & Y &\stackrel{\text{def}}{=} \mathrm{E}\{y\}, \\ z &\stackrel{\text{def}}{=} \mathcal{J}\left(\hat{\alpha}_G^{(c)}; \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right)\right), & Z &\stackrel{\text{def}}{=} \mathrm{E}\{z\} \end{aligned}$$

$$c = X - Y, \ \ d = X - 2Z + Y.$$

Then, within a first-order approximation of (2.15), the variance of $\hat{\alpha}_G^{(f)}$ is given by:

$$\begin{aligned} \mathrm{Var}\{\hat{\alpha}_G^{(f)}\} = 2\Delta^2 \Big[&\left(\frac{d-c}{d^2}\right)^2 \mathrm{Var}\{x\} + \left(\frac{d+c}{d^2}\right)^2 \mathrm{Var}\{y\} + \left(\frac{2c}{d^2}\right)^2 \mathrm{Var}\{z\} \\ &- \left(\frac{d^2-c^2}{d^4}\right) \mathrm{Cov}\{x,y\} + \left(\frac{2dc+2c^2}{d^4}\right) \mathrm{Cov}\{z,y\} + \left(\frac{2dc-2c^2}{d^4}\right) \mathrm{Cov}\{x,z\} \Big] \end{aligned} \tag{2.24}$$

As far as the ML estimator is concerned, the same results apply, where the appropriate objective function as in (2.22) is considered instead of $\mathcal{J}\left(\alpha; \mathbf{W}_0\left(\hat{\alpha}_G^{(c)}\right)\right)$

The mean values X, Y, Z and the variance and covariances of x, y, z of both the objective functions have been analytically evaluated; the details of the derivation are reported in Apps. B and C.

2.6 Final Remarks

The approach here proposed is definitely general and can be applied whenever a statistic can be found on which the *estimandum* acts as a location parameter, providing a CAN estimation at a low computational complexity, also when the pdf of the data is either unknown or analytically intractable. With respect to the general GMM framework, the shift nature of the parameter helps

1. to solve the, generally non-convex, *unweighted* GMM minimization problem in an extremely computational efficient way, resorting to a DFT-based approach

2. to perform the inversion of the variance-covariance matrix with a significant complexity reduction, via a simple columns and row cyclic shift of the previously evaluated covariance matrix for $\alpha = 0$

3. to perform, with a reduced complexity, the direct maximization of the log-likelihood function, given that the observations are multinomially distributed.

In the following chapters we will apply these general principles to the case study of blind phase acquisition for QAM constellations.

The reasons why this case study has been chosen are the followings. First, although the ML cost function is straightforwardly written in the case of known gain, there is no exact way to maximize it and approximate methods have been found only for high or low SNR; then, exhaustive search techniques or suitably initialized gradient-based search techniques should be implemented, at a very high computational cost. Besides, in case of unknown gain, proper evaluation of the ML cost function requires a suitable averaging over all the possible gain determinations; otherwise, MLE for known gain has to be applied after a gain control stage, and this worsens the ML estimator performance. Second, although several suboptimal estimators have been envisaged that differently trade-off efficiency versus computational complexity, the finding of a gain-control-free high SNR near efficient estimator is still open, especially on high cardinality constellations.

Chapter 3

GMM-based Phase Offset Estimation: an Image Processing Approach

3.1 Introduction

The phase acquisition problem is common to many band-pass digital communication systems. In fact, even when the communication channel is ideal, *i.e.* it solely attenuates and delays the transmitted signal, the random carrier phase rotation between local oscillators generates an unknown phase-offset that must be recovered resorting to trained or blind methods, depending on the application requirements.

Many blind phase estimation algorithms have been devised by exploiting Higher Order Statistics of the received signals. In (23), an estimator based on the evaluation of the average of the fourth-order power of the received signal samples has been developed by modification of the estimator described in (24). It is worth noting that the estimator in (23) does not need any gain control. In (25), it is described an estimator approximating the Maximum-Likelihood (ML) estimator for low signal-to-noise ratio values; in (26), this estimator has been shown to be equivalent to the fourth-power estimator of (23). Several improvements to (23) have been investigated, resorting to eighth-order statistics (27), to data aided estimation (28), and to ML estimation for not equiprobable symbols (29). In (30),

a phase histogram based estimator has been derived according to a high SNR approximation of the ML criterion. The estimator requires a preliminary hard-decision stage and, for that, it needs gain control already performed; moreover, as shown in (12), its accuracy presents a constellation dependent penalty with respect to Cramér Rao Bound (CRB).

Much research efforts have been spent to strive the estimation accuracy loss due to the fact that, due to the finite number of measurements, the constellation points are observed with relative frequencies that differ from their nominal probabilities. This effect definitely limits the estimator performance in high SNR since it appears even when additive disturbances are not present. Even though this circumstance does not substantially affect the accuracy of estimators operating in tracking mode, where decision-directed based estimation techniques are usually employed, it definitely limits the accuracy of estimators operating in acquisition mode. In the sequel, we will refer to this effect as *assortment-noise* .

To reduce the assortment-noise, the received data can be suitably selected by certain nonlinear transformations (26; 31), but this selection results in severe accuracy loss for high cardinality constellations. In (21), it is described a generalized form of low-SNR ML approximation that is based on a suitable, constellation dependent, nonlinear pre-processing stage, which also requires knowledge of both the signal gain and the noise power. The estimator accuracy presents the same slope as CRB, but, since the ML approximation substantially holds only for small order constellations, it exhibits a performance loss for increasing constellation cardinality. This loss is mainly due to the nonlinear pre-processing of the received signal samples, which counteracts the above said finite sample effects by filtering out those samples judged unreliable for the estimation, thus reducing the overall number of samples used in the estimation.

However, existing carrier acquisition techniques typically require a preliminary gain adjustment stage. On the other hand, although assuming perfect gain knowledge it is possible to approach CRB in high SNR, as shown by some results presented in (12), the accuracy of gain estimation does affect both the estimation performance and the computational complexity of the estimation process.

Here we present a novel blind phase-offset estimator for general QAM signals. The estimator stems form the observation that, under the presence of a phase

offset θ, the constellation diagram of the received samples, results counterclockwise shifted of θ with respect to the reference case of zero phase offset. The information about the phase offset is then encoded in the constellation diagram of the received samples, and can be evaluated by measuring the shift of each constellation point with respect to its original position in the diagram. In this sense, the radial position of the constellation points carries no information about the phase offset. The dependence on the magnitude of the received samples can be then suppressed by a suitable saturation of the magnitude variable.

In detail we estimate this shift by measuring the cyclic shift of the pdf of the phase of the received signal with respect to the reference pdf of the received samples with zero phase offset.[1] The received signal phase pdf is estimated through a canonical histogram procedure, and estimation of the cyclic phase shift is conducted according to well known optimal matched-filtering and time-delay estimation techniques,[2] by calculating the cross-correlation between the phase histogram measured on the received data and a reference phase pdf analytically evaluated within the zero phase-offset hypothesis. The location of the maximum of such cross-correlation is taken as the estimate of the unknown phase-offset. Moreover, due to the periodicity modulo 2π of the phase, the cross-correlation is a cyclic one and efficient Fast Fourier Transform (FFT) based techniques can be fruitfully employed.

The cyclic cross-correlation between the measured phase histogram and a reference, analytically evaluated, pdf constitutes a concrete novelty with respect to previous histogram based estimation procedures presented in (30), which, in addition, require a preliminary gain adjustment and magnitude hard decision stage. The magnitude hard decision stage directly originates from the high SNR approximation of the ML criterion, and, quoting from (30), it is needed to select "only the (constellation) symbols whose amplitude most closely matches the amplitude of the received data". In our approach, the reference phase pdf is

[1] In fact, the counterclockwise shift on the constellation diagrams, translates on the pdf of the received signal phase as a cyclic shift by θ.

[2] See (35) for a derivation of optimal discrete matched-filters, and (15) and reference therein included for a discussion of optimal discrete-time techniques for time-delay estimation, respectively.

first analytically evaluated within the zero phase-offset hypothesis, and then the phase histogram calculation is carried out without any preliminary gain adjustment and magnitude hard decision. When evaluated around its maximum, the cyclic cross-correlation achieves the matching of the received data, accounted for by the weighed phase histogram, and the constellation symbols, accounted for by the reference phase pdf.

Actually, we present the estimation procedure in an extended version that considers the tomographic projection of the bidimensional (magnitude/phase) pdf of suitable nonlinear transformations of the received data. This, in turn, results in a magnitude weighed phase histogram and in an increased overall estimation accuracy, as shown by theoretical analysis and numerical simulations here performed to assess the estimator performance.

In summary, the novel phase-offset estimator is characterized by a low computational complexity and it does not need the knowledge of both the signal gain and the noise power, but only of the SNR.[1]

Since the said tomographic projection of the magnitude/phase pdf will play a key role during the whole of this work, we will refer to it as the *Constellation Phase Signature* (CPS).[2] Throughout this Chapter, we will analytically evaluate the mean, the variances/covariances of the sample CPS. We will also evaluate the analytical form of the ideal CPS, thus extending a result obtained in (30), and we will show that the sample CPS is an unbiased and consistent estimator of the ideal CPS. Moreover, we will also conduct the asymptotical performance analysis of the related, cross-correlation based, phase-offset estimator.

To assess the estimator performance, we will show a set of numerical simulation results that illustrates the comparison between the accuracies of the CPS based estimator and of some selected existing blind estimators. In particular, as the SNR increases, the comparison shows that the slope of the curve of the CPS based estimator variance and the slope of CRB, calculated in the case of

[1] We will show that only a very rough estimate SNR estimate is needed.

[2] In principle, the estimator does require the knowledge of the signal constellation; however, we remark that the CPS itself exhibits relevant constellation dependent features that can be exploited for constellation detection (classification). Thus, unless differently stated, we will assume in the following that the constellation is known, either *a priori*, or through a preliminary constellation detection stage.

unknown gain, as detailed in Chapter 5, are the same. Let us also say that accuracy comparison has been performed assuming, when needed, perfect knowledge of the signal gain for competitor estimators, while our estimator is gain control free.

This estimation technique can be correctly recast in the GMM-based framework, and, moreover, given the cyclic shift nature of the *estimandum*, it is a valid case study for the GMM estimation procedure introduced in the pervious chapter. We will show that the CPS-based estimator can be interpreted as an application of the *unweighted* GMM estimation procedure and we will extend the estimation rule in the sense of the *best* GMM estimation, *i.e.* the optimally weighted estimator that minimizes the estimation variance. Moreover, we carry out the asymptotical performance analysis for both the unweighted and the best estimator. We report several simulation results that assess the performance analysis and show that a significant accuracy improvement is attained by the *best* GMM estimator at medium-high Signal-to-Noise Ratio (SNR) values, where the weighted estimator approaches the CRB. Furthermore, we show that the values of the CPS, under a particular setting, can be recognized to represent frequencies of recurrence of the received samples phase, so that the reduced complexity MLE, as detailed in Section 2.4.2, can be evaluated. Simulation results show that, for high SNR, the CRB is attained with extremely low complexity.

3.2 Discrete-Time Signal Model

Let us consider a QAM based communication link, and let $S[n]$ be the n-th transmitted symbol drawn from a power normalized M-ary constellation $\mathcal{A} = \{S_0, .., S_{M-1}\}$. At the receiver side, after front-end signal processing, a complex low-pass version of the received signal is available for sampling, and let $X[n]$ be the samples of the received signal extracted at symbol rate. We will assume the following analytical model of the signal samples $X[n]$:

$$X[n] = G_c\, e^{j\theta}\, S[n] + W[n] \tag{3.1}$$

where G_c is the unknown overall gain, θ is the unknown phase-offset, and $W[n]$ is a realization of a circularly complex Gaussian stationary noise process, statistically

independent of $S[n]$. The signal-to-noise ratio (SNR) is defined as $\mathrm{SNR} \stackrel{\mathrm{def}}{=} G_c^2/\sigma_W^2$, being $\sigma_W^2 \stackrel{\mathrm{def}}{=} \mathrm{E}\{|W[n]|^2\}$ the noise variance.

3.3 Image Processing Approach to Phase-Offset Estimation

According to (23; 25; 30; 32), we retain a zero frequency-offset assumption, in order to address here only the estimation of the constant[1] phase-offset θ given a sample of N consecutive observations $X[n],\ n = 0, \ldots, N-1$. The estimation is carried out within an inherent $\pi/2$ ambiguity interval due to the quadrant constellation symmetry. Once the phase has been recovered and compensated, the sequence of received samples is still rotated by an unknown value $k\pi/2,\ k = 0, \ldots, 3$; this ambiguity can be recovered by training-based procedures or bypassed if differential phase encoding or other suitable symbol coding is adopted (see for instance (34)).

Following the guidelines in (21), let us consider the following nonlinear functionof the received signal samples $X[n]$:

$$Y[n] = |X[n]|^P \cdot e^{j4 \cdot \arg X[n]} \tag{3.2}$$

Let us remark that measurements whose phases differ by $\pi/2$ are indistinguishable in (3.2).

Nonlinear transformations folding the received samples by assigning the phase argument $4 \cdot \arg X[n]$ are explicitly considered in (21; 26) and implicitly applied in (30), where the received signal samples are projected in the first quadrant to take into account the constellation symmetry. In (32), a particular case of (3.2) is exploited, namely the fourth-order power of the measurements ($P = 4$).

[1]When the phase estimator is applied in presence of an unknown frequency offset, this latter must be estimated and compensated before the phase estimation stage. Thus, a degradation of the phase estimator performance is expected. In fact, for a large number of observation, the CRB for unbiased phase-offset estimators is approximately one quarter of the same CRB calculated in presence of a frequency offset nuisance parameter (12). A CPS based frequency estimator is presented in the next Chapter.

Generally speaking, the nonlinearity (3.2) tends to cluster the measurements in particular angular windows of the complex plane, in a fashion that depends on the transmitted constellation. In fact, let us observe that, after the nonlinearity (3.2), measurements drawn from a specific constellation exhibit a typical *fingerprint*, *i.e.* a particular clustered distribution of $Y[n]$.

In Fig.3.1 we report the constellation diagrams of the received samples, and the nonlinearly transformed samples, along with the corresponding constellation points obtained in the reference case of zero phase offset and SNR $\to \infty$. From these diagrams it is clear the effect of both the phase offset, which counterclockwise shifts the clusters of $Y[n]$ by 4θ and the gain G_c, which spreads the clusters over the radial direction. The information about the phase offset is directly encoded in the phase location of each cluster in Fig.3.1, while the radial location carries no information about the phase offset. To suppress the influence of the gain nuisance parameter, we need to saturate the radial dependency on the magnitude of the clusters of which $Y[n]$ is constituted.

This can be obtained by considering a magnitude weighted tomographic projection of the bidimensional pdf of the magnitude/phase of the received samples. What we expect after this operation, is a monodimensional function constituted by a number of pulses whose number and location depend on the clusters in Fig.3.1 (and hence on the constellation), and whose shape depends on the noise distribution. In this domain the counterclockwise shift produced on the cluster by the presence of phase offset translates in a cyclic shift of 4θ with respect to the zero phase offset reference case.

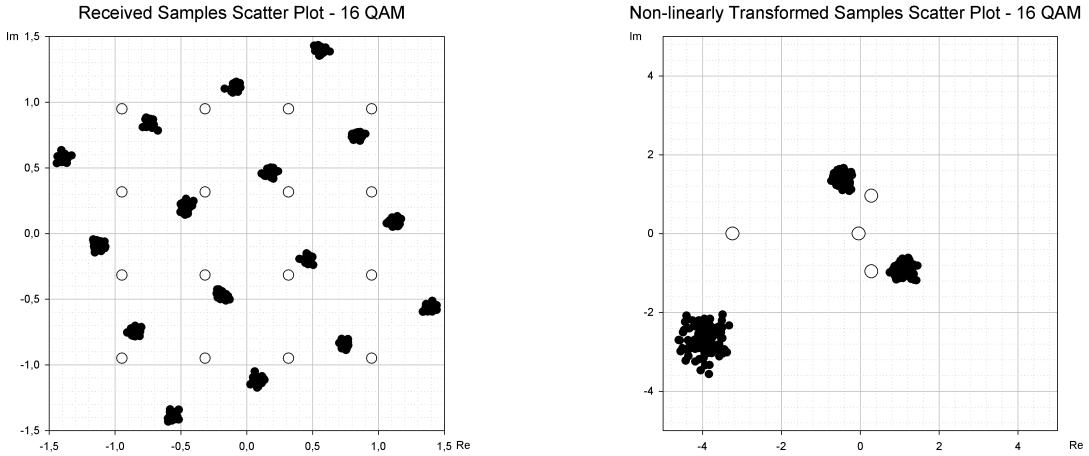

Figure 3.1: *Received samples $X[n]$ (left) and nonlinearly transformed samples $Y[n]$ (right), 16-QAM. The white dots refer to the reference noise free case with zero phase offset, i.e. $\theta = 0$.*

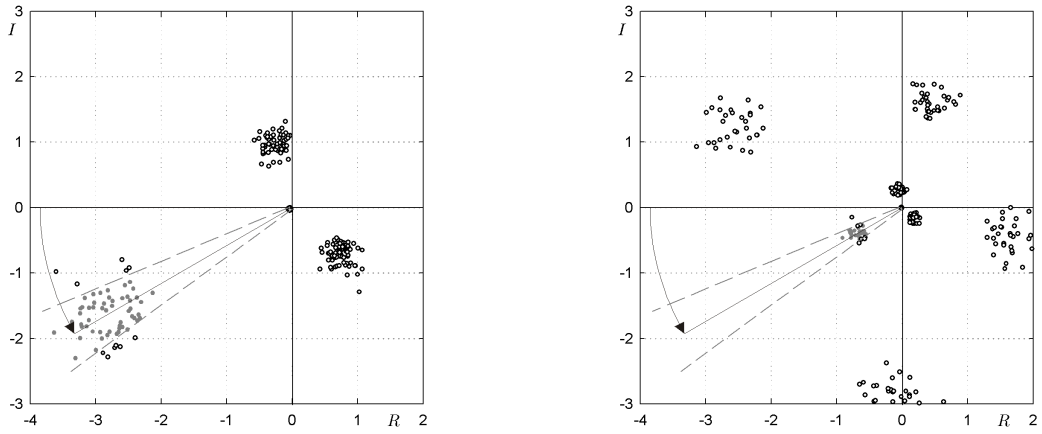

Figure 3.2: *Area of accumulation of nonlinear transformed received signal samples, $P = 4$,16-QAM and 32-QAM constellation, noisy case.*

Let us drop the time index n for the sake of readability. In the noise free case, the pdf of the random variable Y is expressed by a suite of Dirac pulses. Let us consider the polar representation $Y = re^{j\varphi}$, so to write, for equiprobable

constellation symbols:

$$p_{R,\Phi}(r,\varphi;\theta) = \frac{1}{M}\sum_{m=0}^{M-1}\delta(r - G^P|S_m|^P)\cdot\delta(\varphi - 4\theta - 4\arg S_m) \tag{3.3}$$

where we have not explicitly taken into account that certain groups of pulses share the same phase location.[1]

The marginal pdf of the phase φ is obtained by saturating (3.3) with respect to the magnitude r:

$$p_{\Phi}(\varphi;\theta) \stackrel{\text{def}}{=} \int_0^{+\infty} p_{R,\Phi}(r,\varphi;\theta)\,dr = \frac{1}{M}\sum_{m=0}^{M-1}\delta(\varphi - 4\theta - 4\arg S_m) \tag{3.4}$$

From (3.4), we see that the phase pdf $p_{\Phi}(\varphi;\theta)$ cyclically shifts of 4θ under a phase-offset θ:

$$p_{\Phi}(\varphi;\theta) = p_{\Phi}(\varphi - 4\theta;0)$$

In presence of additive noise, the Dirac pulses in (3.4) become wider pulses whose shape depends on the SNR and the noise pdf,[2] but the cyclic shift due to the phase-offset θ still applies.

Now, let us consider a magnitude weighed version of (3.4), namely:

$$\begin{aligned} g_{\Phi}^{(\mathcal{A},\theta,P)}(\varphi) &\stackrel{\text{def}}{=} \int_0^{+\infty} r\cdot p_{R,\Phi}(r,\varphi;\theta)\,dr \\ &= \frac{1}{M}\sum_{m=0}^{M-1} G_c^P|S_m|^P\delta(\varphi - 4\theta - 4\arg S_m) \end{aligned} \tag{3.5}$$

In (3.5) we have performed a tomographic projection of the magnitude weighed phase pdf $r\,p_{R,\Phi}(r,\varphi;\theta)$; we will refer to the function $g_{\Phi}^{(\mathcal{A},\theta,P)}(\varphi)$ as the Magnitude

[1] From this expression we recognize that θ is a location parameter for the pdf family $p_{R,\Phi}(r,\varphi;\theta)$ according to the condition in (2.1).

[2] For SNR values high enough so to retain the approximation

$$Y[n] \simeq G^P e^{j4\theta}|S[n]|^P\left(1 + \frac{\Re\left\{W[n]e^{-jP\cdot\arg S[n]}\right\}}{|S[n]|^P}\right) \cdot e^{j4\arg S[n]}\exp\left(j4\frac{\Im\left\{W[n]e^{-j4\cdot\arg S[n]}\right\}}{|S[n]|^4}\right)$$

the shape of the Dirac pulses becomes equal to the pdf of the imaginary component of the noise $W[n]$.

Weighed Tomographic Projection (MWTP) of the magnitude/phase bidimensional pdf $p_{R,\Phi}(r,\varphi;\theta)$. From (3.5), we recognize that also the MWTP $g_{\Phi}^{(\mathcal{A},\theta,P)}(\psi_k)$ cyclically shifts of 4θ under a phase-offset θ, and that, in presence of additive noise, also the Dirac pulses in (3.5) become wider pulses whose shape depends on the SNR and the noise pdf; the analytical evaluation of (3.5) is reported in App. E.

Since $g_{\Phi}^{(\mathcal{A},\theta,P)}(\varphi)$ substantially behaves like an ordinary pdf, its estimation requires subdividing the phase interval $[0,2\pi)$ in K intervals of amplitude $2\pi/K$:

$$I_K^{(k)} \stackrel{\text{def}}{=} \left[k\cdot\frac{2\pi}{K},(k+1)\cdot\frac{2\pi}{K}\right), \quad \text{for } k=0,\ldots,K-1 \tag{3.6}$$

Now, let us consider the normalized area of $g_{\Phi}^{(\mathcal{A},\theta,P)}(\varphi)$ in the k-th phase interval:

$$f^{(\mathcal{A},\theta,P)}(\psi_k) \stackrel{\text{def}}{=} \frac{K}{2\pi}\cdot\int_{2\pi k/K}^{2\pi(k+1)/K} g_{\Phi}^{(\mathcal{A},\theta,P)}(\varphi)\; d\varphi \tag{3.7}$$

where $\psi_k \stackrel{\text{def}}{=} 2\pi k/K$ denotes the reference phase of the k-th phase interval. Note that in the limit $K\to\infty$ it results:

$$\lim_{K\to\infty} f^{(\mathcal{A},\theta,P)}(\psi_k) = g_{\Phi}^{(\mathcal{A},\theta,P)}(\psi_k)$$

From (3.7), we see that also the function $f^{(\mathcal{A},\theta,P)}(\psi_k)$ cyclically shifts of 4θ under a phase-offset θ; since it is constellation dependent, in the sequel we will refer to it as the *Constellation Phase Signature* (CPS).

Note that the gain G_c influences the CPS in a multiplicative fashion, so that the cyclic shift is entirely and solely due to the phase-offset θ. In this sense we attained the aim of suppressing the misleading due to the presence of the gain nuisance parameter G_c. Hence, once a sample estimate of the CPS has been obtained, we can rephrase the phase-offset estimation problem as a gain control free, cyclic shift estimation problem, which, in turn, can be solved according to well established, optimal matched-filtering and time-delay estimation techniques.

3.4 Histogram Estimation of the Constellation Phase Signature

To estimate the CPS, we follow the guidelines of the canonical histogram procedure, *i.e.* we consider the accumulation of values $|Y[n]|$ in angular windows of suitable width. Fig.3.2 illustrates an example of a generic angular window

on which the samples $Y[n]$ are accumulated, in the case of noisy 16-QAM and 32-QAM constellations and $P = 4$.

Dropping out the time-index n for the sake of readability, let us define the indicator function $d_K^{(k)}(Y)$ to specify that $\arg Y$ belongs to the k-th phase interval:

$$d_K^{(k)}(Y) \stackrel{\text{def}}{=} \begin{cases} 1 & \arg Y \in I_K^{(k)} \\ 0 & \text{otherwise} \end{cases}$$

Then, we consider the following CPS estimator at the angle $\psi_k \stackrel{\text{def}}{=} 2\pi k/K$, for $k = 0, \ldots, K-1$:

$$\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k) = \frac{K}{2\pi} \cdot \frac{1}{N} \sum_{n=0}^{N-1} |Y[n]| \cdot d_K^{(k)}(Y[n]) \tag{3.8}$$

where we recall that the dependence on the parameter P is implicit in the definition of $Y[n]$.

According to (3.8), the sample CPS $\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)$ is calculated as the sample average of $|Y[n]|$ in the k-th phase interval $I_K^{(k)}, k = 0 \ldots K-1$. For $P = 0$, the sample CPS is a canonical phase histogram calculated on N realizations of the random variable $Y[n]$,[1] while for $P > 0$ the magnitude weight $|Y[n]|$ properly takes into account the tomographic nature of the CPS. As we will show in the sequel, this weighing results in improved final estimation accuracy.

The sample CPS of a M-QAM received signal affected by a phase-offset θ, observed in absence of noise and in the ideal condition of relative frequencies of constellation points exactly equal to their nominal probability, *e.g.* $1/M$ for equiprobable constellation points, reduces to:

$$\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k) = \frac{K}{2\pi} \cdot \frac{1}{M} \sum_{m=0}^{M-1} |G_c S_m|^P d_K^{(k)}\left(e^{j4(\theta + \arg S_m)}\right) \tag{3.9}$$

[1] In (30), the phase histogram is directly calculated on the received data $X[n]$, operating modulo $\pi/2$ to take into account the four-quadrant constellation symmetry. Note that this is equivalent to operate on $4 \arg X[n]$. The introduction of the nonlinear transformed observation $Y[n]$ in (3.2) (and its pdf) allows for more compact definition of the CPS, and the subsequent statistical analysis can be conducted in a more readable form.

Fig.3.3 illustrates the theoretical MWTP [1] for different QAM constellations (SNR = 35dB) in the case of zero phase-offset. It is worth noting that each MWTP shows a number of peaks equal to the number of clusters formed by the samples $Y[n]$ originated by constellation points differing by $\pi/2$.

Fig.3.4, instead, plots the noise-free sample CPS $\hat{f}^{(\mathcal{A},0,P)}(\psi_k)$ for a number of QAM constellations, in the case of unit gain, zero phase-offset and ideal condition of absence of unequal occurrences of the signal constellation points (assortment-noise free). As expected, in ideal condition, the noise-free sample CPS of a given constellation is constituted by a discrete number of samples, whose locations correspond to the various phases of the fourth power of the constellation symbols. Noteworthily, in ideal condition of absence of both assortment-noise and additive noise, from (3.9) we see that also the sample CPS cyclically shifts by 4θ under a phase-offset θ. In presence of both assortment and additive noise, the differences between the noisy sample CPS and the noise-free one, can be appreciated in the plots on the right of Fig.3.4. In particular, the spreading of the pulses is due to the additive noise and the pulse shapes resemble the form of the pdf of the imaginary component of the additive noise. Moreover, due to the assortment-noise, in the sample CPS we cannot observe the ideal property that pair of pulses symmetrically located with respect the CPS maximum possess the same "area".

Other than the magnitude weighing, the main difference with the phase histogram procedure described in (30) is that here we directly accumulate the phase of the received data, while in (30) it is rather accumulated the phase difference taken with respect to the constellation symbol having the closest magnitude. This magnitude hard decision does need a preliminary gain control stage, avoided in our approach. Quoting from (30), the magnitude hard decision is needed to select "only the (constellation) symbols whose amplitude most closely matches the amplitude of the received data", according to a high SNR ML approximation; then, the phase-offset is estimated as the mode of the so built phase histogram. Despite the high SNR ML approximation, in (12) it is shown that the estimation accuracy of such histogram based phase-offset estimator does not approach CRB.

[1] As we stated in the previous section, for K large, the ideal CPS tends to the MWTP. From now on we will always assume K to be such that we can consider $f^{(\mathcal{A},\theta,P)}(\psi_k) \simeq g_{\Phi}^{(\mathcal{A},\theta,P)}(\varphi)$.

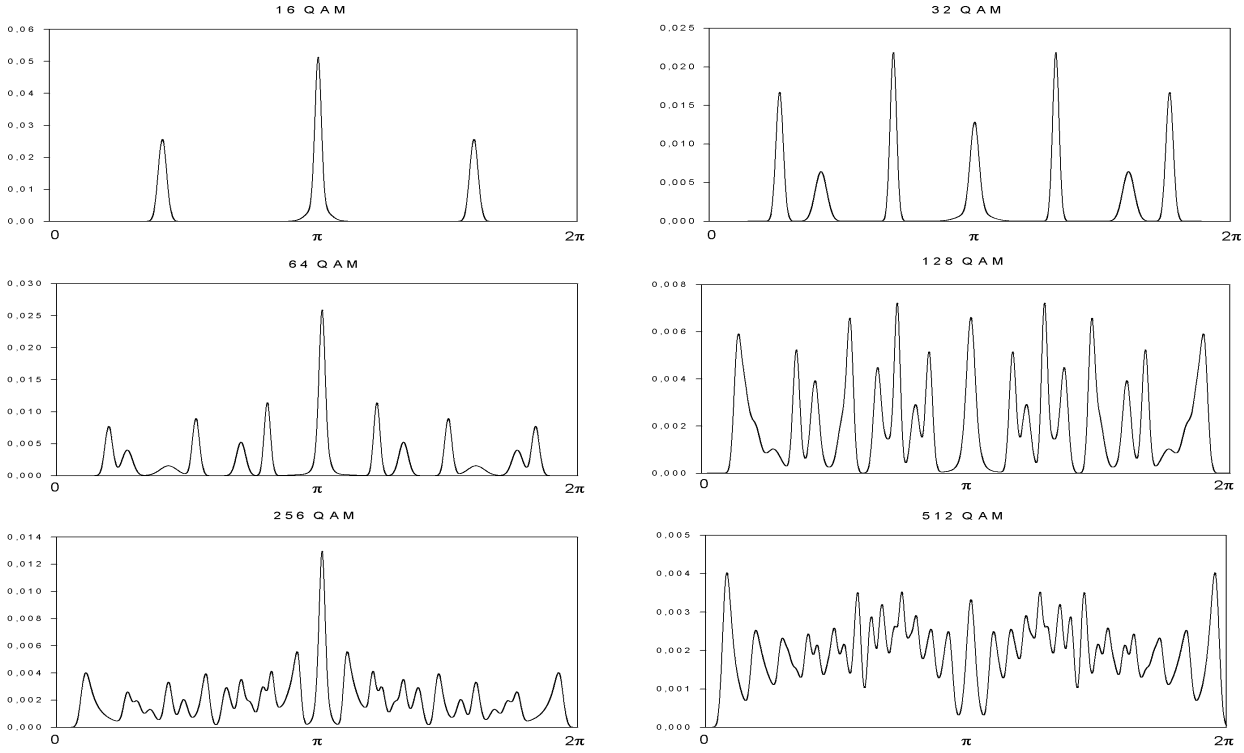

Figure 3.3: *Theoretical MWTP* $g_{\Phi}^{(\mathcal{A},0,P)}(\varphi)$*, for different QAM constellations (SNR* $= 35$dB*).*

As we will see in the sequel, here the phase-offset is rather estimated as the cyclic shift of the CPS, and this is accomplished searching for the maximum of the cyclic cross-correlation between the sample CPS and a reference MWTP calculated under the zero phase-offset hypotheses. Thus, at lags around its maximum, the cyclic cross-correlation carries out, in a joint fashion, the symbol matching required by the high SNR ML approximation between received data and constellation symbols.

Now, after having introduced the CPS, outlined its properties, and described its estimation through a canonical histogram procedure, we are ready to devise a novel phase-offset estimator that directly measures the cyclic shift of the sample CPS, according to optimal matched-filtering and time-delay estimation techniques.

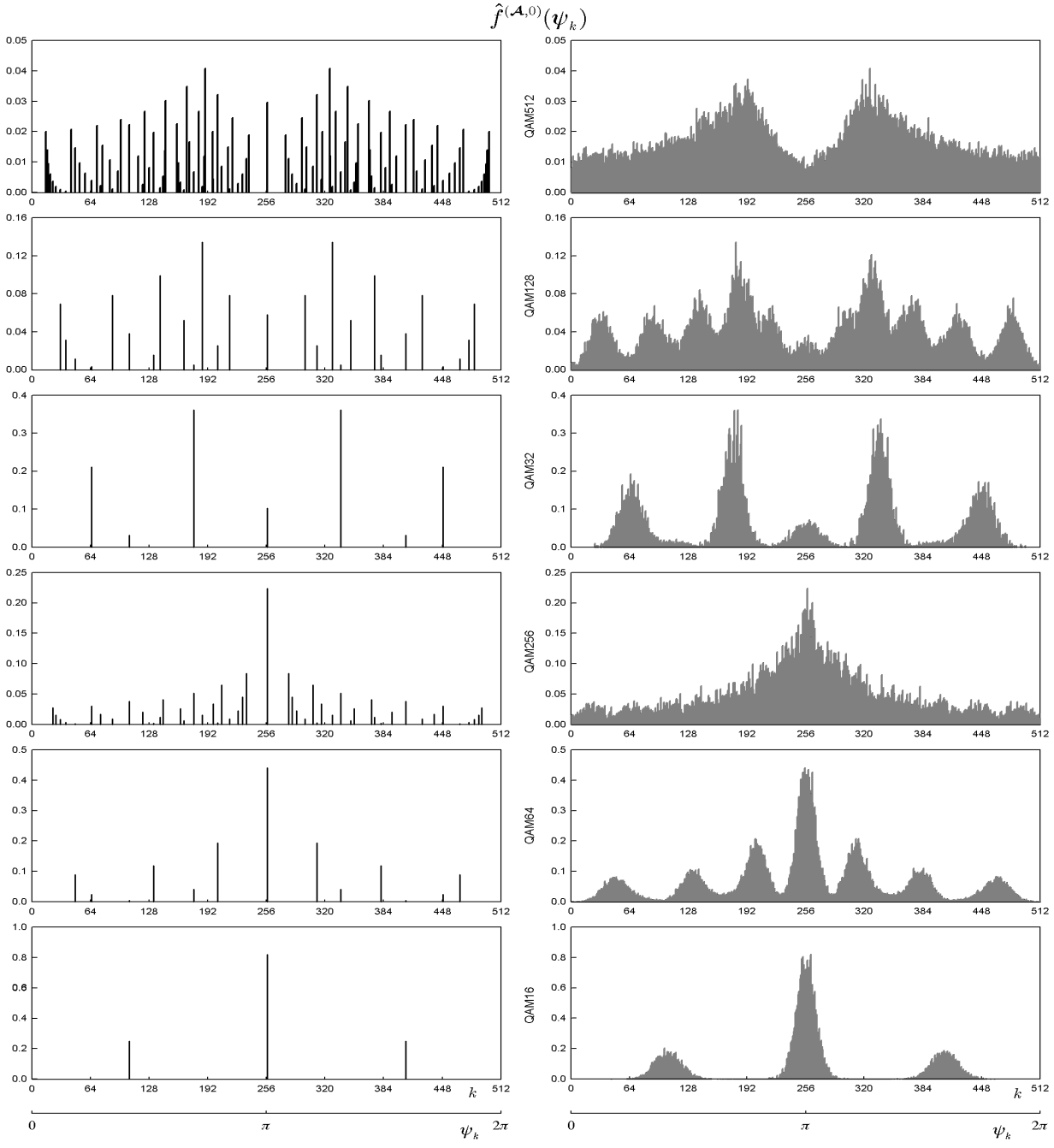

Figure 3.4: *Noise free (left) and noisy (right, SNR* = 23dB*) sample Constellation Phase Signatures* $\hat{f}^{(\mathcal{A},0,P)}(\psi_k)$ *vs. the phase intervals index* k*, for different QAM constellations,* $(\theta = 0, P = 4, K = 512)$*. For the sake of readability, it is also reported the axis of the reference phase of the* k*-th interval, i.e.* $\psi_k = 2\pi k/K$*.*

3.5 Constellation Phase Signature Based Phase-Offset Estimation

As discussed in Sect.3.4, in absence of additive noise and assortment-noise, the sample CPS is constituted by a number of discrete samples of known location and amplitude, and the phase-offset θ results in a CPS cyclic shift of 4θ. Hence, in this ideal transmission environment, an estimate of the phase-offset θ would be obtained estimating the cyclic displacement between the sample CPS $\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)$, as defined in (3.8) and the ideal sample CPS calculated for zero phase-offset, as given in (3.9).

Actually, with reference to (3.1), both the additive noise and the assortment-noise randomize and smooth the sample CPS. Specifically, while the assortment-noise modifies the relative amplitudes of the sample CPS pulses leaving their locations unaltered, the additive noise spreads the locations of the received samples around the constellation points, thus spreading the CPS pulses. Hence, to devise an accurate phase estimator we need to substitute the ideal CPS (3.9) with a suitable template that properly takes into account the additive noise effect.

As shown in App. E, in the limit $K \to \infty$, the sample CPS is an unbiased estimator of the MWTP $g_{\Phi}^{(\mathcal{A},\theta,P)}(\psi)$:

$$\lim_{K\to\infty} \mathrm{E}\left\{\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)\right\} = g_{\Phi}^{(\mathcal{A},\theta,P)}(\psi_k) \tag{3.10}$$

The consistency of the sample CPS is proven in Sect.3.6.1, where its variances - covariances are analyzed in detail.

We recall that $g_{\Phi}^{(\mathcal{A},\theta,P)}(\psi)$ equals the zero phase-offset $g_{\Phi}^{(\mathcal{A},0,P)}(\psi)$ shifted by 4θ along the ψ axis:

$$g_{\Phi}^{(\mathcal{A},\theta,P)}(\psi) = g_{\Phi}^{(\mathcal{A},0,P)}(\psi - 4\theta)$$

Hence, the phase-offset θ can be estimated by evaluating the cyclic shift between the sample CPS $\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)$ and its expectation MWTP $g_{\Phi}^{(\mathcal{A},0,P)}(\psi)$, being this latter analytically calculable. The detailed analytical evaluation of the zero phase-offset MWTP $g_{\Phi}^{(\mathcal{A},0,P)}(\psi)$ is found in App. E. Here, without loss of generality, we report the final expression for the reference case $G_c = 1$, in which the noise

variance σ_W^2 assumes the physical meaning of Noise-to-Signal Ratio:[1]

$$g_{\Phi}^{(\mathcal{A},0,P)}(\psi) = \frac{4}{M}\sum_{m=0}^{M-1}\frac{1}{\pi\,\sigma_W^2}\exp\left(-\frac{\gamma_m^2\sin^2(\psi/4-\phi_m)}{\sigma_W^2}\right) \cdot\sum_{p=0}^{P+1}\binom{P+1}{p}\left(\gamma_m\cos(\psi/4-\phi_m)\right)^{P+1-p}U^{(p)}(\psi,\gamma_m,\phi_m,\sigma_W) \tag{3.11}$$

where $S_m = \gamma_m e^{j\phi_m}$ is the polar representation of the $m-$th constellation point, and

$$U^{(p)}(\psi,\gamma_m,\phi_m,\sigma_W) \stackrel{\text{def}}{=} \int_{-\gamma_m\cos(\psi/4-\phi_m)}^{\infty}\gamma^p\exp\left(-\frac{\gamma^2}{\sigma_W^2}\right)d\gamma \tag{3.12}$$

We have analytically evaluated the integrals $U^{(p)}(\psi,\gamma_m,\phi_m,\sigma_W)$, and results for $p=0,\ldots,5$ are reported in App. E.

Let us point out that, since the sample CPS is an unbiased and consistent estimator of the MWTP, the right part of Fig.3.4 also depicts the basic behaviours of these latter.

Now, the estimation of the cyclic phase-offset shift between the sample CPS $\hat{f}^{(\mathcal{A},\theta,P)}(\psi)$ and the reference MWTP $g_{\Phi}^{(\mathcal{A},0,P)}(\psi)$, having analytically evaluated this latter in (3.11), can be conducted through optimal filtering matched to MWTP $g_{\Phi}^{(\mathcal{A},0,P)}(\psi)$, and, precisely, by locating the abscissa of the maximum of the matched filter output. Following the guidelines of (15), we consider a *coarse/fine* two-stage estimation procedure. A coarse phase-offset estimate $\hat{\theta}_c$ is first obtained as follows:

$$\begin{aligned}\hat{C}[k] &\stackrel{\text{def}}{=} \hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)\otimes g_{\Phi}^{(\mathcal{A},0,P)}(\psi_k)\\ \hat{k}_0 &= \arg\max_k\left\{\hat{C}[k]\right\}\\ \hat{\theta}_c &= \frac{2\pi}{K}\cdot\frac{\hat{k}_0}{4}\end{aligned} \tag{3.13}$$

where the symbol $\otimes$ denotes cyclic cross-correlation with respect to the phase index k.[2]

[1] The result in (3.11) for $P=0$ is found also in Appendix B of (30).

[2] Since the additive noise that affects $\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)$ is colored, optimal matched filtering of

Let us observe that, according to (3.11), the evaluation of the zero phase-offset MWTP $g_{\Phi}^{(\mathcal{A},0,P)}(\psi_k)$ requires the knowledge of the SNR G^2/σ_W. In Sect.3.7, devoted to the experimental results and performance comparison, we will discuss the sensitivity of the CPS based estimation accuracy to perturbations of SNR, and we will show that our phase-offset estimator is sufficiently robust to SNR mismatch errors, because this mismatch mainly affects the width, but not the locations, of the pulses of the MWTP.

The cyclic cross-correlation (3.13) can be carried out resorting to classical FFT based techniques. The sample CPS $\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)$ is evaluated using the nonlinearly transformed received signal samples, while the FFT of the reference MWTP $g_{\Phi}^{(\mathcal{A},0,P)}(\psi_k)$ can be pre-calculated and stored.

Since the number of phase intervals K measures the CPS sampling rate, in principle it should be chosen as a function of the expected CPS bandwidth, which in turn depends on both the constellation and the SNR; on the other hand, K can be fixed according to the desired estimator computational complexity. In Sect.4.6 we will see that, for SNR values of practical interest, the choice $K=512$ constitutes an acceptable trade-off between the overall computational complexity and the CPS based estimation accuracy.

Let us remark that, once the CPS is sampled on a discrete number of phase intervals K, the coarse estimate $\hat{\theta}_c$ is limited at the resolution $2\pi/K$. As indicated in (15), the coarse phase estimation can be refined by interpolating the cross-correlation values around the index k_0; the parabolic approximation of the cyclic cross-correlation function yields the following fine phase-offset estimate:

$$\hat{\theta}_f = \hat{\theta}_c + \frac{1}{4}\frac{2\pi}{K}\frac{1}{2}\frac{\hat{C}[k_0+1]-\hat{C}[k_0-1]}{\hat{C}[k_0+1]+\hat{C}[k_0-1]-2\hat{C}[k_0]} \tag{3.14}$$

3.5.1 Final Remarks

Apart the magnitude weighing, in comparing the CPS based estimation technique with the histogram based technique presented in (30), the first important

the sample CPS $\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)$ requires a noise whitening prefiltering before the cyclic cross-correlation (3.13). Anyway, we have not observed any misdetection in all the numerical experiments performed using (3.13); for this reason, we have preferred to present an estimation procedure with reduced computational complexity.

difference to be noted is that we do not make any preliminary gain control and hard decision, operations that in (30) are performed according to a high SNR ML approximation to, namely, "select constellation symbols whose amplitude closely matches the amplitude of received data". Thus, a (symbol matched) differential phase histogram is calculated therein, whose mode is taken as the phase-offset estimate; here the phase histogram is directly constructed on the received data, and, rather than the mode, its cyclic shift is taken as the phase-offset estimate. The cyclic shift is measured by the abscissa of the maximum of the cyclic cross-correlation between the sample CPS $\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)$ and the reference MWTP $g_{\Phi}^{(\mathcal{A},0,P)}(\psi_k)$; the optimality of this estimation procedure is well established within the framework of optimal matched-filtering and time-delay estimation.

Moreover, in our approach, the symbol matching required by the high SNR ML approximation is implicitly performed at the maximum of the cyclic cross-correlation. In fact, the maximum of the cyclic cross-correlation occurs when every peak of the sample CPS, corresponding to a subset of received data, is aligned with a specific, corresponding MWTP peak. The symbol matching is achieved because this MWTP peak is formed with those, and only those, constellation symbols carried by the considered received data subset.

Noteworthily, the matching is jointly performed over all the received data, and computational efficiency is assured by using FFT based cyclic cross-correlation computation.

Even though the resolution of the matching between symbols and received data is limited by the number of intervals K selected to construct the sample CPS, the theoretical performance analysis conducted in Sect.3.6, confirmed by numerical simulation results shown in Sect.3.7, indicates that the accuracy of our estimator approaches CRB; as reported in (12), this does not happen for the histogram based estimator of (30). The accuracy improvement obtained by our estimator is essentially due to the fact that all the received data corresponding to the various constellation points are jointly exploited in the cyclic cross-correlation based phase-offset estimation, and reduced noise sensitivity is observed with respect to the estimation that in (30) considers the location of the mode of the phase histogram.

Besides, it is worth mentioning that our gain control free technique obtains significant performance improvement in medium/highSNR also with respect to the optimal nonlinear least-square estimation technique described in (21), in spite of the fact that this latter operates with perfect gain knowledge.

3.6 Performance analysis

In this Section, the performance analysis of the CPS based phase-offset estimator is outlined. In Subsect.3.6.1 we first evaluate the variances/covariances of the sample CPS $\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)$. Then, in Subsect.3.6.2 we evaluate the variances/covariances of the sample cyclic cross-correlation function $\hat{C}[k]$. Finally, in Subsect.3.6.3, we complete the analysis by evaluating the bias and the variance of the fine phase-offset estimator $\hat{\theta}_f$ as functions of the variances/covariances of $\hat{C}[k]$.

3.6.1 Variance and Covariances of the Sample CPS

Here, we report only the main results concerning mean, variances and covariances of the sample CPS $\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)$; analytical details are found in App. E and App F. We will also prove that, in the limit $K \to \infty$, the sample CPS is an unbiased and consistent estimator of the MWTP.

In fact, bearing (3.8) in mind, the mean of the sample CPS is:

$$\mathrm{E}\left\{\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)\right\} = \mathrm{E}\left\{|Y| \cdot d_K^{(k)}(Y)\right\}$$

and in App. E we have shown that, in the limit $K \to \infty$:

$$\mathrm{E}\left\{\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)\right\} = g_{\Phi}^{(\mathcal{A},\theta,P)}(\psi_k)$$

The asymptotic variances/covariances of the sample CPS are:

$$N\,\mathrm{Var}\left\{\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)\right\} = \mathrm{E}\left\{|Y|^2 \cdot d_K^{(k)}(Y)\right\} - \left(\mathrm{E}\left\{|Y| \cdot d_K^{(k)}(Y)\right\}\right)^2$$

$$N\,\mathrm{Cov}\left\{\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k), \hat{f}^{(\mathcal{A},\theta,P)}(\psi_l)\right\} \underset{(k\neq l)}{=} -\mathrm{E}\left\{|Y| \cdot d_K^{(k)}(Y)\right\} \mathrm{E}\left\{|Y| \cdot d_K^{(l)}(Y)\right\}$$

and in App. F we have shown that, in the limit $K \to \infty$:

$$N \cdot \mathrm{Var}\left\{\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)\right\} = g_{\Phi}^{(\mathcal{A},\theta,2P)}(\psi_k) - \left(g_{\Phi}^{(\mathcal{A},\theta,P)}(\psi_k)\right)^2$$

$$N \cdot \mathrm{Cov}\left\{\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k), \hat{f}^{(\mathcal{A},\theta,P)}(\psi_l)\right\} \underset{(k\neq l)}{=} -g_{\Phi}^{(\mathcal{A},\theta,P)}(\psi_k) \cdot g_{\Phi}^{(\mathcal{A},\theta,P)}(\psi_l)$$

3.6.2 Variances and Covariances of the Sample Cross-Correlation

In order to evaluate the variances and the covariances of the sample cyclic cross-correlation function $\hat{C}[k]$, here we write the sample CPS $\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)$ as the sum of its expected value $g_{\Phi}^{(\mathcal{A},\theta,P)}(\psi_k)$ and a zero mean, estimation error $N[k]$:

$$N[k] = \hat{f}^{(\mathcal{A},\theta,P)}(\psi_k) - g_{\Phi}^{(\mathcal{A},\theta,P)}(\psi_k) \tag{3.15}$$

Hence, we have

$$\mathrm{Var}\{N[k]\} = \mathrm{Var}\left\{\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)\right\}$$

and

$$\mathrm{Cov}\{N[k], N[l]\} = \mathrm{Cov}\left\{\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k), \hat{f}^{(\mathcal{A},\theta,P)}(\psi_l)\right\}$$

Let us observe that, in computing the cyclic cross-correlation function, the estimation error $N[k]$ that affects the sample CPS is coloured by the reference zero phase-offset MWTP $g_{\Phi}^{(\mathcal{A},0,P)}(\psi_k)$:

$$\hat{C}[k] = \hat{f}^{(\mathcal{A},\theta,P)}(\psi_k) \otimes g_{\Phi}^{(\mathcal{A},0,P)}(\psi_k) = C[k] + E[k] \tag{3.16}$$

where $C[k] \overset{\mathrm{def}}{=} g_{\Phi}^{(\mathcal{A},\theta,P)}(\psi_k) \otimes g_{\Phi}^{(\mathcal{A},0,P)}(\psi_k)$ and

$$E[k] = N[k] \otimes g_{\Phi}^{(\mathcal{A},0,P)}(\psi_k) \tag{3.17}$$

Let us decompose the phase-offset as $\theta = 2\pi k_0/K + \theta_f$, being k_0 the index of the maximum of $C[k]$ with $|\theta_f| < \pi/K$, and let us introduce the following compact notation:

$$\mathbf{N} \overset{\mathrm{def}}{=} \big[N[k_0] \cdots N[K-1] N[0] \cdots N[k_0-1]\big]^{\mathrm{T}}$$

and, for $l = 0, \ldots, K-1$:

$$\mathbf{G}_l \stackrel{\text{def}}{=} \Big[\underbrace{g_{\Phi}^{(\mathcal{A},0,P)}(\psi_l) \cdots g_{\Phi}^{(\mathcal{A},0,P)}(\psi_{K-1})}_{K-1-l} \; \underbrace{g_{\Phi}^{(\mathcal{A},0,P)}(\psi_0) \cdots g_{\Phi}^{(\mathcal{A},0,P)}(\psi_{l-1})}_{l} \Big]^{\mathrm{T}}$$

Let us observe that $g_{\Phi}^{(\mathcal{A},\theta,P)}(\psi_{k+k_0}) = g_{\Phi}^{(\mathcal{A},\theta_f,P)}(\psi_k)$, and that, in the limit $K \to \infty$, we can set $\theta_f = 0$. Hence, the asymptotical covariance matrix of the error $\mathbf{N}$ takes the following form:

$$\mathbf{K}_{\mathrm{N}} \stackrel{\text{def}}{=} N\mathrm{E}\left\{\mathbf{N}\mathbf{N}^{\mathrm{T}}\right\} = \mathrm{diag}\left[g_{\Phi}^{(\mathcal{A},0,2P)}(\psi_0), g_{\Phi}^{(\mathcal{A},0,2P)}(\psi_1), \cdots, g_{\Phi}^{(\mathcal{A},0,2P)}(\psi_{K-1})\right] - \mathbf{G}_0\mathbf{G}_0^{\mathrm{T}}$$

Hence, the expectation and the asymptotical second order moments of the error $E[k+k_0]$ are given by:

$$\mathrm{E}\{E[k+k_0]\} = 0 \tag{3.18}$$

$$r_{EE}[k,l] \stackrel{\text{def}}{=} N \cdot \mathrm{E}\{E[k+k_0]E[l+k_0]\} = \mathbf{G}_k^{\mathrm{T}} \cdot \mathbf{K}_{\mathrm{N}} \cdot \mathbf{G}_l \tag{3.19}$$

3.6.3 Variance of the Fine Phase-Offset Estimator

In evaluating the accuracy of the estimator (3.14), we observe that two error components appear. The first component occurs when the coarse estimate $\hat{\theta}_c$ is not correct due to $\hat{k}_0 \neq k_0$. The second error component affects the fine estimate $\hat{\theta_f}$ and it is due to the finite sample size and to the misfit of the parabolic approximation around its maximum, and definitely limits the phase-offset estimator accuracy. As far as the coarse estimate is concerned, numerical simulations have confirmed $\hat{k}_0 = k_0$ for a large range of SNR values, and thus we discuss only the bias and the variance of $\hat{\theta}_f$. Following the approach indicated in (15), by setting

$$c = C[k_0+1] - C[k_0-1] \qquad ; \qquad d = C[k_0+1] - 2C[k_0] + C[k_0-1]$$

and resorting to the following first-order approximation of (3.14):

$$\hat{\theta}_f - \hat{\theta}_c \approx \frac{\pi}{4L}\left(\frac{c}{d} + \frac{d-c}{d^2}E[k_0+1] - \frac{d+c}{d^2}E[k_0-1] - \frac{2c}{d^2}E[k_0]\right)$$

we have obtained the following result:

$$\begin{aligned}
&\mathrm{aVar}(\hat{\theta}_f) \stackrel{\text{def}}{=} N \cdot \mathrm{Var}(\hat{\theta}_f) = \\
&\frac{\pi^2}{16K^2}\Big[\left(\frac{d-c}{d^2}\right)^2 r_{EE}[1,1] + \left(\frac{d+c}{d^2}\right)^2 r_{EE}[-1,-1] + \left(\frac{2c}{d^2}\right)^2 r_{EE}[0,0] \\
&- \left(\frac{d^2-c^2}{d^4}\right) r_{EE}[1,-1] + \left(\frac{2dc+2c^2}{d^4}\right) r_{EE}[0,-1] + \left(\frac{2dc-2c^2}{d^4}\right) r_{EE}[1,0]\Big]
\end{aligned}$$

As far as the bias of $\hat{\theta}_f$ is concerned, we observe that, since the ideal cyclic cross-correlation $C[k]$ is not exactly parabolic around its maximum value located at $k = k_0$, a systematic bias due to the parabolic misfit is present (see (15) for a discussion on the parabolic misfit bias). However, for the range of K of interest the values $C[k_0 - 1], C[k_0], C[k_0 + 1]$ are sufficiently close so that the bias due to the parabolic approximation is negligible in evaluating the estimation Mean Square Error (MSE). For this reason, in the performance comparison reported in Sect.4.6, the theoretical analysis results will be expressed in terms of the error standard deviation $\sqrt{\mathrm{Var}\,(\hat{\theta}_f)}$, while the experimental results will be expressed in terms of Root Mean Square Error (RMSE).

3.7 Numerical Experiments and Performance Comparison

Here, we present numerical results aimed at validating the theoretical performance analysis of Sect.3.6, and at comparing the accuracy of the estimator described in Sect.3.5 with selected state-of-the-art estimators. Each experiment consists of 2000 MonteCarlo trials and in each trial the phase-offset has been drawn from a uniform distribution in $[-\pi, \pi)$.

The analytical asymptotical error standard deviation and the experimental RMSE of the phase estimator versus SNR are shown in Figs. 3.5-3.10, which illustrate the results for the estimator $\hat{\theta}_c$ (SIG-COARSE), the estimator $\hat{\theta}_f$ (SIG-FINE) ($P = 1$), along with the results of the theoretical analysis carried out in Sect.3.6 and with CRB for unknown gain evaluated following the guidelines in Chapter 5. In detail, Figs. 3.5, 3.7 and 3.9 correspond to square constellations whereas Figs. 3.6, 3.8 and 3.10 correspond to cross constellations. The graphs plot the experimental results obtained with sample size $N = 500$ for square constellations and $N = 2000$ for cross constellations.

From Figs.3.5-3.10 we see an excellent agreement between the analytical performance evaluated in Sect.3.6 and the numerical results, either for square or cross constellations, showing that the large sample approximation underlying the

asymptotical analysis holds true for such realistic values of N or larger. Moreover, the asymptotical RMSE of the SIG-FINE estimator versus SNR presents the same slope of the CRB.

For comparison purposes, also results pertaining to other selected estimators are reported, namely the Nonlinear Least Square (WA03) estimator (21), the Concentration Ellipse Orientation (CEO) estimator (32), and the fourth-order (CW99) estimator (23) for square constellations or the eighth-order (CW01) estimator (27) for cross constellations. In particular, the WA03 estimator is obtained by evaluating the angular coordinate of the sample average of a suitable nonlinear transformation of the received signal samples; it is constellation dependent and it requires knowledge of both SNR and gain. It must be remarked that perfect knowledge of only the SNR has been assumed in the implementation of our CPS based estimator, while perfect knowledge of both the gain and the SNR has been assumed in the implementation of the WA03 estimator. The CEO estimator is obtained by evaluating the orientation of the concentration ellipse of the normal distribution fitting the distribution of the fourth power of the received signal samples. The CW99 and CW01 estimators are obtained by evaluating the angular coordinate of the fourth order and of the eighth order sample averages, respectively. The CEO, CW99 and CW01 are constellation independent and do not require knowledge of the gain and of the SNR. Performance comparison with the estimators WA03, CEO, CW99 and CW01 shows that the CPS based estimator outperforms CEO, CW99, CW01 and WA03 estimators, apart the case of 16-QAM where it equals the WA03 estimator. It is worth noting also that WA03 estimator approximates ML estimator for small-order constellations.

For dense constellations, the CPS based estimator presents a significant accuracy improvement over all estimators for medium to high values of SNR; at high SNR the accuracy improvement is more pronounced since the CPS estimator better counteracts the assortment-noise effects. In medium-high SNR, the accuracy improvement ranges from a few dB for small-order constellations up to $10 \div 15$dB for large order cross constellations. The performance improvement with respect to WA03 estimator is mainly due to the fact that the CPS based estimator employs all the received samples, while WA03 estimator discards an increasing number of samples as the constellation cardinality increases. The

performance improvement with respect to the blind estimators CEO, CW99 and CW01 is due to the fact that, when the constellation is known, the phase-rotation of the measurement clusters in the complex plane is better captured by the CPS cyclic cross-correlation mechanism, which, in turn, performs the symbol matching indicated by the high SNR ML approximation.

Different values of K have been tested, and we can observe that a reduction of K to 256 or an increase up to 1024 does not modify the CPS based estimator accuracy at medium-low SNR, and only a slight accuracy loss is observed for cross constellations at high SNR. This is exemplified in Tab.3.1 and 3.2, where the numerical results concerning with the asymptotical RMSE at SNR = 40dB are reported for square and cross constellation, respectively.

$\sqrt{N}$ RMSE @ 40dB	$K = 256$	$K = 512$	$K = 1024$
16-QAM	0.58	0.56	0.56
64-QAM	0.63	0.61	0.62
256-QAM	0.87	0.84	0.84

Table 3.1: Asymptotical RMSE of $\hat{\theta}_f$ at SNR$=$40dB for different values of K (square constellations, $N=500$, $P=1$, 2000 MonteCarlo trials).

$\sqrt{N}$ RMSE @ 40dB	$K = 256$	$K = 512$	$K = 1024$
32-QAM	0.62	0.54	0.54
128-QAM	0.66	0.63	0.59
512-QAM	0.94	0.85	0.81

Table 3.2: Asymptotical RMSE of $\hat{\theta}_f$ at SNR$=$40dB for different values of K (cross constellations, $N=2000$, $P=1$, 2000 MonteCarlo trials).

The impact of the magnitude power P on the CPS based estimation accuracy is shown in Figs.3.11 and 3.12, where the analytical asymptotical RMSE is plotted for $P=1$ (dotted line), $P=2$ (dashed line), and $P=3$ (continuous line), for square and cross constellations, respectively; we see that high values of P improve the

estimator accuracy at low-SNR values, without increasing the overall computational complexity when a look-up table is employed to evaluate the magnitude power function.

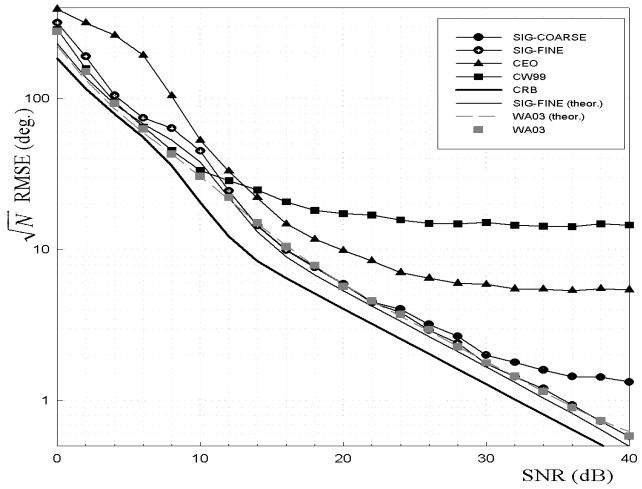

Figure 3.5: *16-QAM constellation: RMSE of various phase estimators vs. SNR; estimator $\hat{\theta}_c$ (SIG-COARSE) and $\hat{\theta}_f$ (SIG-FINE), CEO estimator, Nonlinear Least Square estimator (WA03), fourth-order Cartwright estimator (CW99), theoretical standard deviation, CRB ($N = 500, K = 512, P = 1,2000$ MonteCarlo trials).*

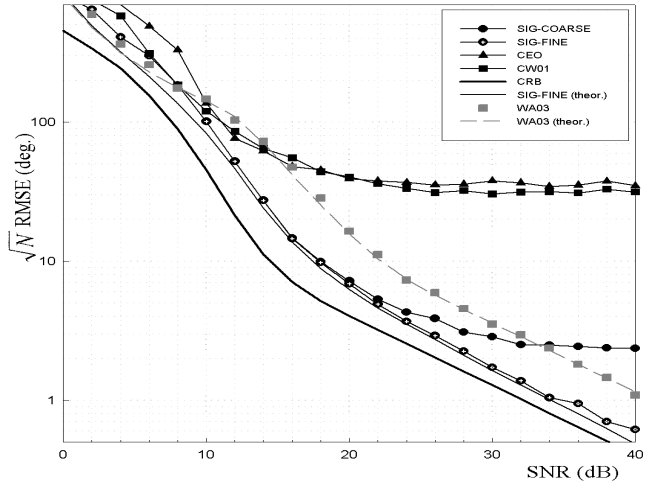

Figure 3.6: *32-QAM constellation: RMSE of various phase estimators vs. SNR; estimator $\hat{\theta}_c$ (SIG-COARSE) and $\hat{\theta}_f$ (SIG-FINE), CEO estimator, Nonlinear Least Square estimator (WA03), eight-order Cartwright estimator (CW01), theoretical standard deviation, CRB ($N = 2000, K = 512, P = 1,2000$ MonteCarlo trials).*

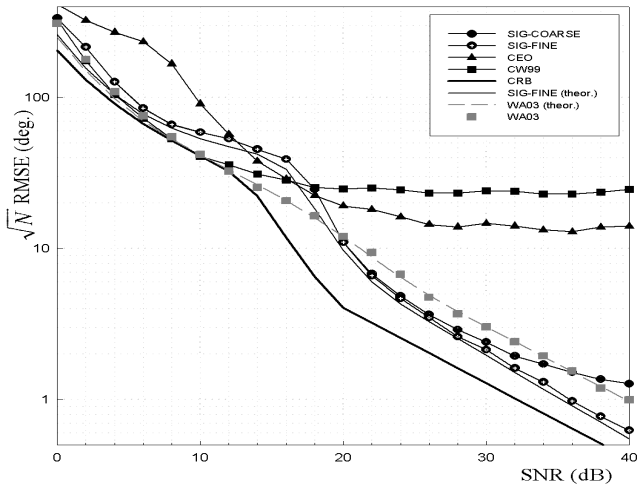

Figure 3.7: *64-QAM constellation: RMSE of various phase estimators vs. SNR; estimator $\hat{\theta}_c$ (SIG-COARSE) and $\hat{\theta}_f$ (SIG-FINE), CEO estimator, Nonlinear Least Square estimator (WA03), fourth-order Cartwright estimator (CW99), theoretical standard deviation, CRB ($N = 500, K = 512, P = 1,2000$ MonteCarlo trials).*

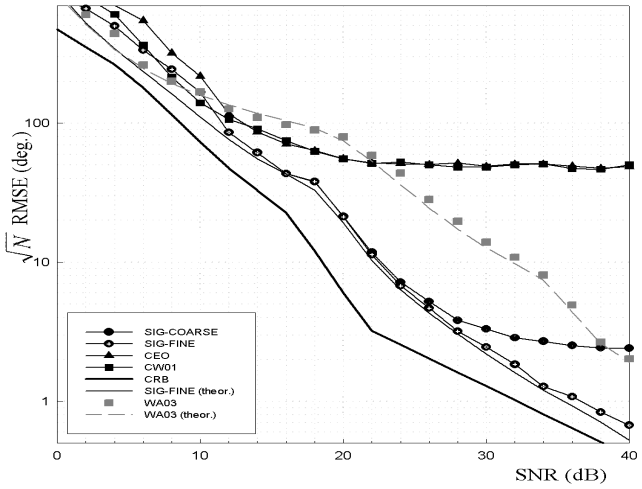

Figure 3.8: *128-QAM constellation: RMSE of various phase estimators vs. SNR; estimator $\hat{\theta}_c$ (SIG-COARSE) and $\hat{\theta}_f$ (SIG-FINE), CEO estimator, Nonlinear Least Square estimator (WA03), eight-order Cartwright estimator (CW01), theoretical standard deviation, CRB ($N = 2000, K = 512, P = 1,2000$ MonteCarlo trials).*

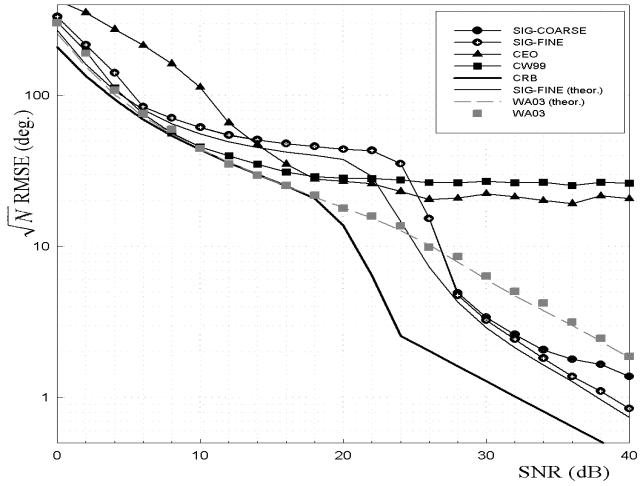

Figure 3.9: *256-QAM constellation: RMSE of various phase estimators vs. SNR; estimator $\hat{\theta}_c$ (SIG-COARSE) and $\hat{\theta}_f$ (SIG-FINE), CEO estimator, Nonlinear Least Square estimator (WA03), fourth-order Cartwright estimator (CW99), theoretical standard deviation, CRB ($N = 500, K = 512, P = 1,2000$ MonteCarlo trials).*

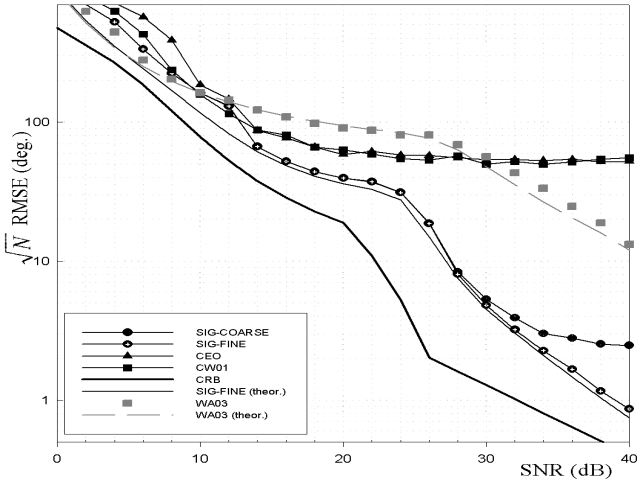

Figure 3.10: *512-QAM constellation: RMSE of various phase estimators vs. SNR; estimator $\hat{\theta}_c$ (SIG-COARSE) and $\hat{\theta}_f$ (SIG-FINE), CEO estimator, Nonlinear Least Square estimator (WA03), eight-order Cartwright estimator (CW01), theoretical standard deviation, CRB ($N = 2000, K = 512, P = 1,2000$ MonteCarlo trials).*

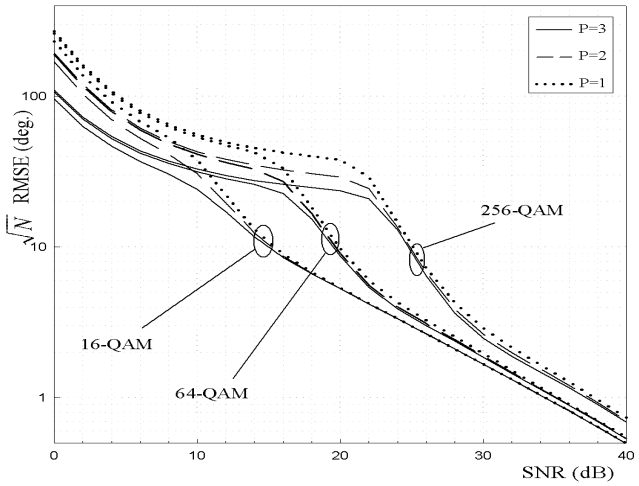

Figure 3.11: *Square constellations: analytical RMSE of the phase estimator* $\hat{\theta}_f$ *vs. SNR for* $P=1$ *(dotted line),* $P=2$ *(dashed line), and* $P=3$ *(continuous line)* ($K=512$, $N=500$, 2000 *MonteCarlo trials).*

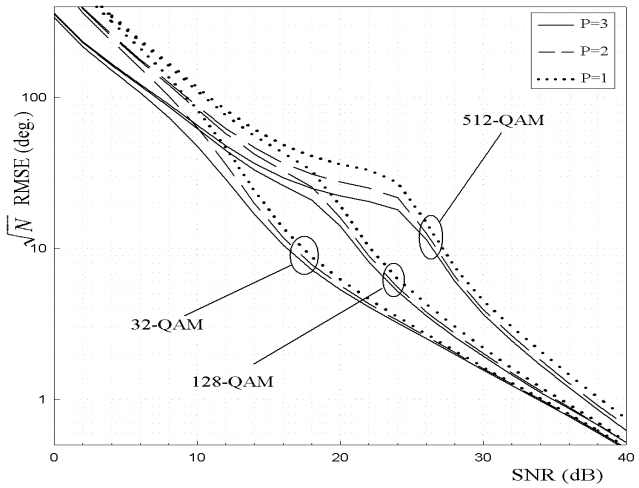

Figure 3.12: *Cross constellations: analytical RMSE of the phase estimator* $\hat{\theta}_f$ *vs. SNR for* $P=1$ *(dotted line),* $P=2$ *(dashed line), and* $P=3$ *(continuous line)* ($K=512$, $N=2000$, 2000 *MonteCarlo trials).*

Finally, we analyze the CPS based estimator performance within the hypothesis that the zero phase-offset reference MWTP $g_{\Phi}^{(\mathcal{A},0,P)}(\psi_k)$ has been calculated in presence of a SNR mismatch. To this purpose, we have evaluated the accuracy of the estimator $\hat{\theta}_f$ (SIG-FINE) under SNR mismatches of $\pm$3dB, $\pm$6dB, and $\pm$10dB and the results are collected in Fig.3.13. The observed estimation robustness with respect SNR mismatches is explained recalling that the phase-offset estimate is obtained from the location of the maximum of the cross-correlation between the sample CPS and the reference MWTP, and that the SNR value is needed only to fix the width of the pulses of the reference MWTP. Hence, relatively small pulse width variations, either resulting in wider or narrower pulses, mainly affect the width of peak around the cross-correlation maximum, while leaving substantially unaltered its location. This explains also why the sensitivity diminishes with the SNR. On the other hand, at high SNR the CPS based estimation accuracy preserves the CRB slope, both for square and cross constellations, even under SNR mismatch of $\pm$10dB.

Several SNR estimators are available from the technical literature, see (36) and the references therein contained for instance. The analysis concerning with the selection of a specific SNR estimator is beyond the scope of this work; nonetheless, we have enriched Fig.3.13 providing also numerical results obtained using an heuristic, rough SNR estimator, better described in App. G, based on the measurement of the bandwidth of the CPS. As it can be seen, the robustness of the CPS based estimator under SNR mismatch is well maintained also using this simple SNR estimator.

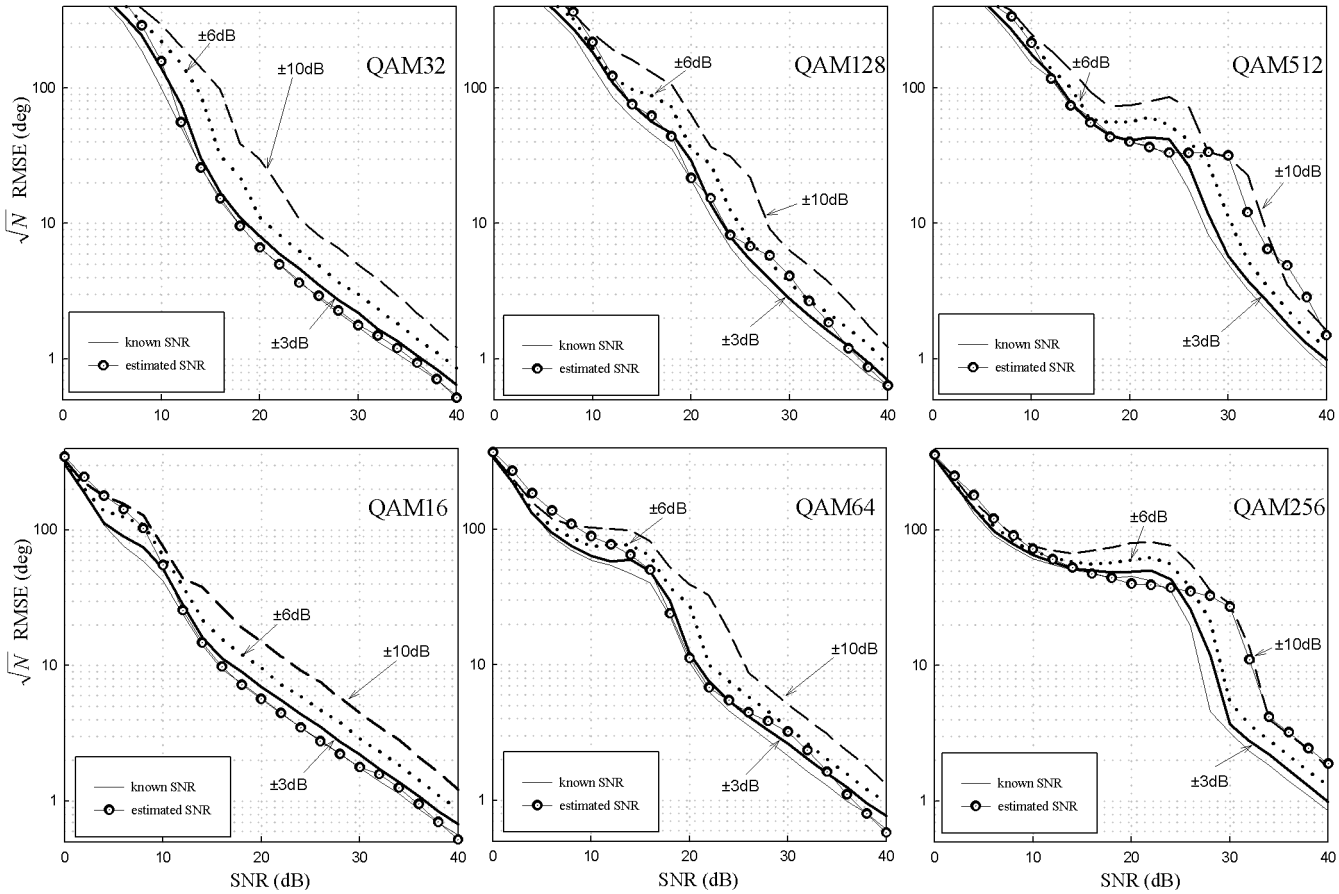

Figure 3.13: *RMSE of the phase estimator $\hat{\theta}_f$ vs. SNR for SNR mismatch equal to* $\pm 3\,\mathrm{dB}$ *(continuous line),* $\pm 6\,\mathrm{dB}$ *(dotted line), and* $\pm 10\,\mathrm{dB}$ *(dashed line). For each doublet of SNR mismatch, the RMSE curve reports the worst measured performance. The RMSE curve obtained using the rough SNR estimator described in App.G is also reported (bulleted line), as well as the RMSE curve obtained with perfectly known SNR (tiny continuous line).*
($K = 512, P = 1,2000$ *MonteCarlo trials,* $N = 500$ *for square constellations,* $N = 2000$ *for cross constellations).*

3.8 GMM based Phase Acquisition for QAM Constellation

3.8.1 Introduction

In the following Section, we will recast the phase offset estimator herein introduced in the GMM framework for cyclic-shift parameter estimation detailed in Chapter 2. We will first prove that the estimator herein presented is an unweighted GMM estimation and then we will extend it in the sense of the *best* GMM estimator. Numerical resuslts show a significant accuracy improvement by the *best* GMM estimator at medium-high Signal-to-Noise Ratio (SNR) values, where the weighted estimator approaches the CRB.

Moreover we will prove that the values of the sample CPS, under a particular setting, can be recognized to be multinomially distributed, and hence a reduced complexity ML estimation will be proposed, according to the guidlines depicted in Subsection 2.4.2.

3.8.2 GMM phase Acquisition

Here we recast the estimator introduced in this Chapter in the framework of the GMM estimation, extending it in the sense of the *best* GMM estimator.

To this end let us observe that the CPS as defined in (3.7) can be compactly rewritten as:

$$f^{(\mathcal{A},\theta,P)}\left(\psi_k\right) = \frac{K}{2\pi}\int_{\psi_k}^{\psi_k+2\pi/K} d\varphi \int_0^{+\infty} r\, p_{R,\Phi}\left(r,\varphi|\theta\right) dr$$

where we can regognize it as an instance of the nonlinear moment defined in (2.2). In the same way, the sample CPS as defined in (3.8), can be easily recognized to be a specification of the weighted histogram in (2.3).

We can then correctly define the observation vector collecting the values of the sample CPS:

$$\hat{\mathbf{f}} = \left[\hat{f}^{(\mathcal{A},\theta,P)}\left(\psi_0\right),\ldots,\hat{f}^{(\mathcal{A},\theta,P)}\left(\psi_{K-1}\right)\right]^{\mathrm{T}} \tag{3.20}$$

and the reference vector collecting the values [1] of the ideal CPS for a generic phase offset ξ:

$$\mathbf{f}(\xi) = \left[f^{(\mathcal{A},\xi,P)}(\psi_0), \ldots, f^{(\mathcal{A},\xi,P)}(\psi_{K-1}) \right]^{\mathrm{T}} \tag{3.21}$$

Now, we can use the following generalized moment:

$$\mathbf{e}(\xi) \stackrel{\mathrm{def}}{=} \hat{\mathbf{f}} - \mathbf{f}(\xi) \tag{3.22}$$

Taking into account the four quadrant phase folding in (3.2), it results:

$$\mathrm{E}_{x|\theta}\{\mathbf{e}(\xi)\} = \mathbf{0} \quad \textit{iff} \quad \xi = 4\theta \tag{3.23}$$

Then, following the approach introduced in Chapter 2, the GMM estimation of the phase-offset can be conducted by minimizing the following objective function:

$$\begin{aligned} \hat{\theta}_G^{(\mathbf{W})} &= \frac{1}{4} \arg\min_{\xi} \mathcal{Q}(\xi; \mathbf{W}) \\ \mathcal{Q}(\xi; \mathbf{W}) &= (\hat{\mathbf{f}} - \mathbf{f}(\xi))^{\mathrm{T}} \mathbf{W} (\hat{\mathbf{f}} - \mathbf{f}(\xi)) \end{aligned} \tag{3.24}$$

In the *unweighted* case, recalling the derivation described in Sect. 2.3 we have:

$$\begin{aligned} \hat{\theta}_G^{(\mathbf{I})} &= \frac{1}{4} \arg\min_{\xi} \mathcal{Q}(\xi; \mathbf{I}) = \frac{1}{4} \arg\max_{\xi} \; \hat{\mathbf{f}}^{\mathrm{T}} \cdot \mathbf{S}(\xi) \cdot \mathbf{f}(0) \\ &= \frac{1}{4} \arg\max_{\xi} \sum_{k=0}^{K-1} \hat{f}^{(\mathcal{A},\theta,P)}(\psi_k) \cdot f^{(\mathcal{A},0,P)}(\psi_k - \xi) \end{aligned} \tag{3.25}$$

Hence, $\hat{\theta}_G^{(\mathbf{I})}$ is obtained as the index of maximum of the cyclic cross-correlation between the sequences collected in the vectors $\hat{\mathbf{f}}$ and $\mathbf{f}(0)$. We recognize that (3.25) is exactly the cost function in (3.13), here re-obtained in the framework of the GMM estimation procedure. Again we can notice that the maximization of (3.25) can be conducted even though the sample CPS is estimated apart an amplitude scale factor, *i.e.* without accomplishing a preliminary gain control stage, and the phase offset estimation has to be properly considered *gain-control-free*.

[1] In the followings we will assume K to be large enough to neglect the aliasing (possibly) present in the ideal CPS sampling, implicitly performed in the evaluation of the vector $\mathbf{f}(\xi)$.

As far as the *best* GMM estimator is concerned, the optimal weight matrix turns out to be: $\mathbf{W}^{(o)} = \mathbf{\Omega}(\theta)^{-1}/N$, being $\mathbf{\Omega}(\theta) \stackrel{\text{def}}{=} \mathrm{E}\{(\hat{\mathbf{f}} - \mathbf{f}(\theta))(\hat{\mathbf{f}} - \mathbf{f}(\theta))^{\mathrm{T}}\}$ the covariance matrix of the measurements, with entries:

$$||\mathbf{\Omega}(\theta)||_{i,j} \stackrel{\text{def}}{=} \mathrm{Cov}\left\{\hat{f}^{(\mathcal{A},\theta,P)}(\psi_i), \hat{f}^{(\mathcal{A},\theta,P)}(\psi_j)\right\} \tag{3.26}$$

The variances and covariances of the estimated CPS, whose expressions are in turn are a specification of the general formulas in App.A, have been evaluated in Subsection 3.6.2.

The optimal estimation is then given by the solution of the following optimization problem: [1]

$$\begin{aligned} \hat{\theta}_G^{(\mathbf{W})} &= \frac{1}{4}\arg\max_{\xi} \mathcal{J}\left(\xi; \mathbf{W}_0\left(\hat{\theta}_G^{(\mathbf{I})}\right)\right) \\ \mathcal{J}\left(\xi; \mathbf{W}_0\left(\hat{\theta}_G^{(\mathbf{I})}\right)\right) &= \hat{\mathbf{f}}^{\mathrm{T}} \cdot \mathbf{W}_0\left(\hat{\theta}_G^{(\mathbf{I})}\right) \cdot \mathbf{f}(\xi) \\ &\quad - \frac{1}{2} \cdot \mathbf{f}(\xi)^{\mathrm{T}} \cdot \mathbf{W}_0\left(\hat{\theta}_G^{(\mathbf{I})}\right) \cdot \mathbf{f}(\xi) \end{aligned} \tag{3.27}$$

The fine estimation is obtained after parabolic interpolation (15) as in (3.28)

$$\begin{aligned} \hat{\theta}_G^{(\mathbf{W})} &= \hat{\theta}_G^{(\mathbf{I})} - \frac{\Delta\theta}{8} \\ &\cdot \frac{\mathcal{J}\left(\hat{\theta}_G^{(\mathbf{I})} + \Delta\theta; \mathbf{W}_0\left(\hat{\theta}_G^{(\mathbf{I})}\right)\right) - \mathcal{J}\left(\hat{\theta}_G^{(\mathbf{I})} - \Delta\theta; \mathbf{W}_0\left(\hat{\theta}_G^{(\mathbf{I})}\right)\right)}{\mathcal{J}\left(\hat{\theta}_G^{(\mathbf{I})} + \Delta\theta; \mathbf{W}_0\left(\hat{\theta}_G^{(\mathbf{I})}\right)\right) - 2\mathcal{J}\left(\hat{\theta}_G^{(\mathbf{I})}; \mathbf{W}_0\left(\hat{\theta}_G^{(\mathbf{I})}\right)\right) + \mathcal{J}\left(\hat{\theta}_G^{(\mathbf{I})} - \Delta\theta; \mathbf{W}_0\left(\hat{\theta}_G^{(\mathbf{I})}\right)\right)} \end{aligned} \tag{3.28}$$

where $\Delta\theta = 2\pi/K$. The estimation form (3.28) is analytically tractable, and it allows to evaluate the asymptotical performance of the optimal GMM based estimation.

[1] In principle, (3.27) requires a preliminary gain estimation stage. Nevertheless, the optimization can be conducted after proper performed normalization of the CPS vectors so as to restore the unitary gain condition,For instance, the CPS vectors can be normalized as:

$$\hat{\mathbf{f}} \implies \frac{\hat{\mathbf{f}}}{\sqrt{\hat{\mathbf{f}}^{\mathrm{T}} \cdot \hat{\mathbf{f}}}} \quad ; \quad \mathbf{f}(\xi) \implies \frac{\mathbf{f}(\xi)}{\sqrt{\mathbf{f}(0)^{\mathrm{T}} \cdot \mathbf{f}(0)}}$$

thus avoiding any preliminary gain control stage. Bearing this in mind, the estimation procedure (3.27) can be conducted in a gain-control-free fashion.

3.8.3 Reduced Complexity ML Phase Acquisition

Now we illustrate how the CPS can be employed to derive an asymptotically efficient phase offset estimation. Let us consider the sample estimate CPS $\hat{f}^{(\mathcal{A},\theta,0)}(\psi_k)$, where the weighting parameter P has been set to zero,[1] that is:

$$\hat{f}^{(\mathcal{A},\theta,0)}(\psi_k) = \frac{K}{2\pi} \cdot \frac{1}{N} \sum_{n=0}^{N-1} d_K^{(k)}(Y[n]) \tag{3.29}$$

with

$$Y[n] = e^{j4 \cdot \arg X[n]} \tag{3.30}$$

We observe that the sample CPS in (3.29) is, in the end, a phase histogram evaluated using N measurements of the nonlinearly transformed received samples $Y[n]$ in (3.30). Hence, let us consider the K mutually exclusive events $\mathcal{E}_k = \left\{ Y[n] : \arg Y[n] \in I_K^{(k)} \right\}, k = 1, \ldots, K-1$, *i.e.* that a received sample's phase, after the transformation in (3.30), lies in the k-th phase interval defined in (3.6); the number of occurrences of $\mathcal{E}_k, k=0,\ldots,K-1$ are given by:

$$\nu_k = \frac{2\pi N}{K} \hat{f}^{(\mathcal{A},\theta,0)}(\psi_k), , k=0,\ldots,K-1$$

The K random variables $\nu_0, \ldots, \nu_{K-1}$ are multinomially distributed:

$$P(\nu_0, \ldots, \nu_{K-1}) = \frac{N!}{\prod_{k=0}^{K-1} \nu_k!} \prod_{k=0}^{K-1} p_k^{\nu_k}$$

where p_k is the probability of the event $\mathcal{E}_k$, *i.e.* $p_k = 2\pi f^{(\mathcal{A},\theta,0)}(\psi_k)/K$.

Stemming from these observations and following the same steps depicted in Sect.2.4.2, the log-likelihood of $\hat{\mathbf{f}}$ can be written as:

$$\begin{aligned} l(\hat{\mathbf{f}}; \mathbf{f}(\xi)) &= A + \frac{2\pi N}{K} \sum_{k=0}^{K-1} \hat{f}_k \ln(f_k(\xi)) \\ A &= \ln N! - \ln \prod_{k=0}^{K-1} \left(\frac{2\pi N}{K} \hat{f}_k \right)! + \frac{2\pi N}{K} \sum_{k=0}^{K-1} \hat{f}_k \ln \frac{2\pi}{K} \end{aligned} \tag{3.31}$$

where $\hat{f}_k(\xi)$'s are the entries of the vector $\hat{\mathbf{f}}$ and $f_k(\xi)$'s are the entries of the vector $\mathbf{f}(\xi)$, respectively defined in (3.20) and (3.21).

[1] Throughout this subsection we will assume P to be set to zero.

Under the hypothesis that the sample CPS represents a sufficient statistic for the estimation of the phase offset under the observation model in (3.1), the ML estimate of the phase offset can be attained by maximizing (3.31) with respect to ξ:

$$\hat{\theta}_{ML} = \frac{1}{4} \arg\max_{\xi} \; l(\hat{\mathbf{f}}; \mathbf{f}\,(\xi)). \tag{3.32}$$

Considering only the terms depending on θ we can simplify (3.32) as follows:

$$\hat{\theta}_{ML} = \frac{1}{4} \arg\max_{\xi} \; \hat{\mathbf{f}}^{\mathrm{T}} \cdot \tilde{\mathbf{f}}(\xi) \tag{3.33}$$

where $\tilde{\mathbf{f}}(\xi) \stackrel{\text{def}}{=} \ln \mathbf{f}(\xi)$.

Then, remembering that ξ affects the ideal CPS as a location parameter, and considering that the normalization factor employed in the definitions of $\hat{\mathbf{f}}$ does not affect the location of the maximum, the solution of the maximization problem is again obtained as the index of maximum of the cyclic cross-correlation between the sample estimate of the CPS $\hat{\mathbf{f}}$ and the logarithm of the analytically evaluated CPS $\tilde{\mathbf{f}}(\xi)$:

$$\begin{aligned} \hat{\theta}_{ML} &= \frac{1}{4} \arg\max_{\xi} \mathcal{C}\,(\xi) \\ \mathcal{C}(\xi) &= \sum_{k=0}^{K-1} \hat{f}^{(\mathcal{A},\theta,0)}\,(\psi_k) \ln f^{(\mathcal{A},0,0)}\,(\psi_k - \xi) \end{aligned} \tag{3.34}$$

It is interesting to notice that (3.34) is very close to the cost function described in Sect.3.5, this latter not comprising the logarithmic nonlinearity over the ideal CPS.

We remark that, as happened in (3.25), the estimation can be conducted even in presence of an amplitude scale factor, the resulting estimator being then properly considered *gain-control-free*.

The estimation is performed exactly as described in Sect. 3.8.2, that is evaluating the DFT of both the sample and the ideal CPS and then obtaining the cyclic cross-correlation after an inverse DFT stage. The only difference is that, in this case, the logarithm of the ideal CPS is considered instead of the CPS itself. Again a parabolic interpolation is to be envisaged to refine the estimate that, after the cross-correlation, is definitely limited by the resolution given by $2\pi/K$.

3.8.4 On the Computational Complexity of the GMM Phase Offset Estimator

A rough comparison of the computational complexity of the GMM and ML estimators is reported in Tab.3.3. For the sake of completeness, since the Nonlinear Least Squares (NLS) estimator described in (21) will be taken into account as a reference for performance comparison, we have reported also its computational complexity. We observe that the MLE computational complexity depends on the product of the constellation cardinality M, of the number of samples N, and of the parameter K that influences the precision $2\pi/K$ of the coarse estimate $2\pi k_c/K$; moreover, for a fair comparison we remark that the MLE requires a preliminary gain estimation stage. Besides the MLE implemented by exhaustive search, we report also the computational complexity of MLE realized by initializing a gradient search algorithm using a preliminary coarse phase estimate. In this case, the complexity does not depend on K but on the overall number of iterations; however, the dependence on the constellation cardinality still makes MLE unfeasible for non trivial constellations. The GMM unweighted estimator approximately requires N operations for the evaluation of the statistics $\hat{\mathbf{f}}$ and $K(1+2\log_2(K))$ for evaluating the cross-correlation in the cost function (3.25). As discussed in **??**, the GMM *best* phase estimator implemented as in (4.19) requires the same operations as the unweighted GMM plus $3K$ operations due to the evaluation of (3.27) in the three points $\alpha = \hat{\alpha}_G^{(c)}$ and $\alpha = \hat{\alpha}_G^{(c)} \pm 2\pi/K$. Finally, we observe that the NLS estimator requires only N operations plus the cost of a preliminary gain estimation.

3.8.5 Numerical Experiments and Performance Comparison

In this Section we illustrate the analytical and numerical performance of the GMM based estimator herein introduced. The experimental conditions are fixed as in Sect.3.7.[1]

[1]For the sake of readibility we briefly recall them: the signal samples are generated according to (3.1) with a sample size $N=512$ for square-constellations (16-64-256 QAM) and $N=2000$ for cross-constellations (32-128-512 QAM); each numerical experiment consists of 1000 Monte

Estimator	Operations
ML–exhaustive search	$M \times N \times K$ + *gain estimation cost*
ML–gradient based	$M \times N \times N_{\text{iterations}}$ + *gain estimation cost*
GMM unweighted	$N + K(1 + 2\log_2 K)$
GMM weighted	$N + K(4 + 2\log_2 K)$
reduced complexity ML	$N + K(1 + 2\log_2 K)$
NLS	N + *gain estimation cost*

Table 3.3: Computational complexity of ML, GMM, reduced complexity ML and NLS phase estimators.

First, in Figs.3.14-3.19, we illustrate the analytical and numerical performance of the unweighted and the *best* GMM estimate, so as to point out the margin improvement that the optimization introduces. The performance at various SNR are illustrated by plotting the results of the theoretical analysis expressed in terms of the normalized standard deviation of the estimation error $\sqrt{N}\cdot\text{StdDev}\{\hat{\theta}_G^{(\mathbf{W})}\}$, and the results of the numerical Monte Carlo simulations expressed in terms of the normalized Root Mean Square Error ($\sqrt{N}\cdot\text{RMSE}$). [1]

Figs.3.14-3.16 refer to square-constellations (16-64-256 QAM) and Figs.3.17-3.19 refer to cross-constellations (32-128-512 QAM). For reference sake, we have also reported the Cramér-Rao Bound (CRB) for unknown gain as evaluated in Chapter 5, and, for fair comparison with state-of-the-art phase estimators, we have plotted in the same Figs. the theoretical and experimental performances of (21); the performance of the unweighted estimator can be correctly recognized to be the same of the CPS estimator as in Sect.3.7.

From these results we observe the appreciable improvement achieved by using the optimized weighting matrix in the weighted GMM estimator with respect to the unweighted GMM one, *i.e.* versus the CPS estimator in Sect.3.5, and w.r.t. the NLS estimator (21). The optimally weighted GMM normalized standard deviation approaches the CRB at medium-to-high SNR.

Carlo trials.

[1] The bias of all the considered estimators is negligible or zero.

The impact of the magnitude power on the optimal GMM estimation accuracy is shown in Figs.3.20 and 3.21, where the analytical asymptotical RMSE is plotted for $P = 1$ (dotted line), $P = 2$ (dashed line), and $P = 3$ (continuous line), for square and cross constellations, respectively; we see that higher values of P improve estimator's performance, especially for low and high SNR, where the estimator approaches the CRB. In the same figures we also reported the theoretical performance of the reduced complexity ML (RCML) estimator introduced in Sect. 3.8.3. We observe that the RCML estimator presents the same performance of the optimal GMM estimator with $P = 1$ with a lower computational complexity approaching the CRB at medium to high SNR values.

From a technical point of view, it is of interest to evaluate the Symbol Error Rate (SER) reduction achieved in correspondence of the phase estimation error variance reduction. At this aim, in Figs. 3.22, 3.23, 3.24 we show the SER obtained after 2000 Montecarlo runs by the GMM unweighted, the GMM weighted and the NLS estimators for the 128-QAM, 256-QAM and 512-QAM constellations, using $N = 512$ samples for 128-QAM and 256-QAM and $N = 2000$ samples for 512-QAM. It is worth noting that the *best* GMM estimator tightly approaches the performance of the Additive White Gaussian Noise (AWGN) channel, outperforming the NLS on almost all the SNR range. For the weighted GMM, the SER curves referring to $K = 512, 256, 128$ are reported. Although the reduction of K reduces the precision of the GMM estimate, the SER is only slightly degraded. Therefore, in the applications the value of K must be selected on the base of the trade-off between the computational complexity issue and the required SER performance.

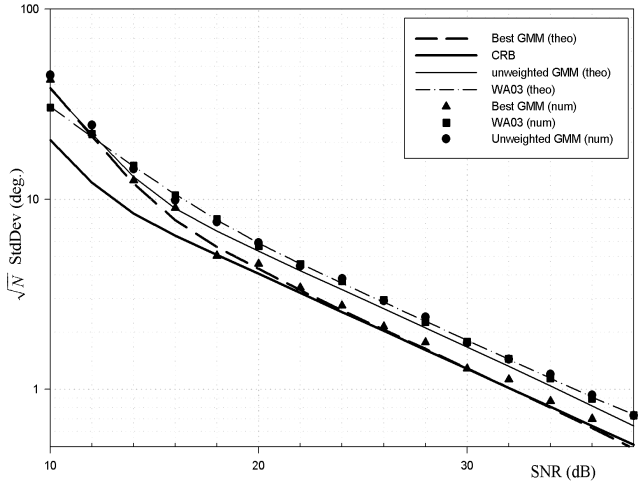

Figure 3.14: *Normalized root mean square phase estimation error* $\sqrt{N} \cdot RMSE$ *vs. SNR for 16-QAM constellation* $(N = 512, K = 512)$*; unweighted GMM estimator (theoretical, solid line gray, and numerical, circles), optimal GMM estimator (theoretical, dashed line gray, and numerical, triangles), NLS estimator in (21) (WA03) (theoretical, dot-dashed line, and numerical, squares). The solid line black represents the CRB.*

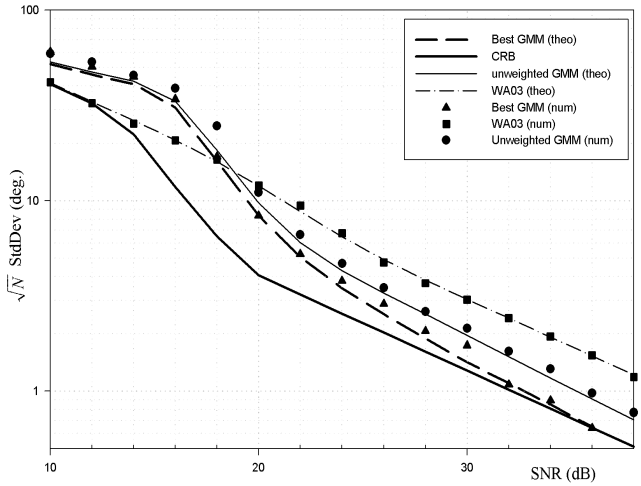

Figure 3.15: *Normalized root mean square phase estimation error* $\sqrt{N} \cdot RMSE$ *vs. SNR for 64-QAM constellation* $(N = 512, K = 512)$*; unweighted GMM estimator (theoretical, solid line gray, and numerical, circles), optimal GMM estimator (theoretical, dashed line gray, and numerical, triangles), NLS estimator in (21) (WA03) (theoretical, dot-dashed line, and numerical, squares). The solid line black represents the CRB.*

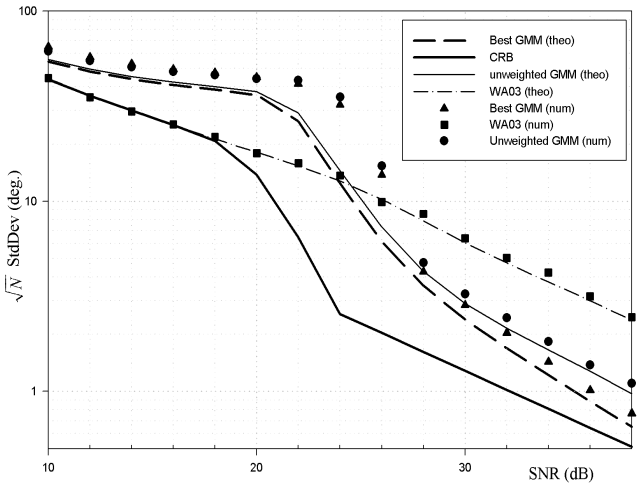

Figure 3.16: *Normalized root mean square phase estimation error* $\sqrt{N} \cdot RMSE$ *vs. SNR for 256-QAM constellation* $(N = 512, K = 512)$*; unweighted GMM estimator (theoretical, solid line gray, and numerical, circles), optimal GMM estimator (theoretical, dashed line gray, and numerical, triangles), NLS estimator in (21) (WA03) (theoretical, dot-dashed line, and numerical, squares). The solid line black represents the CRB.*

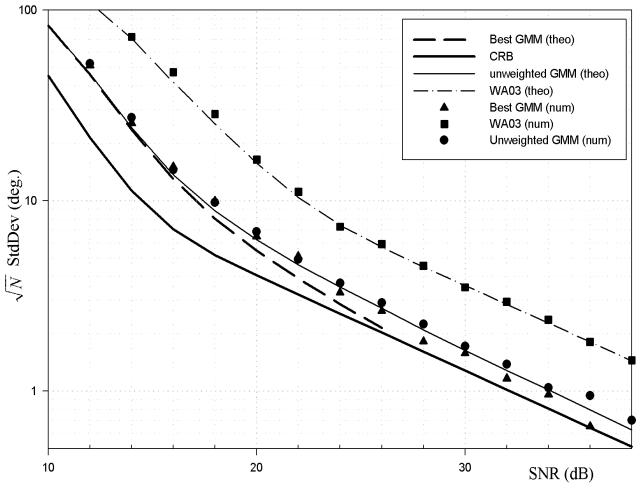

Figure 3.17: *Normalized root mean square phase estimation error* $\sqrt{N} \cdot RMSE$ *vs. SNR for 32-QAM constellation* $(N = 512, K = 2000)$*; unweighted GMM estimator (theoretical, solid line gray, and numerical, circles), optimal GMM estimator (theoretical, dashed line gray, and numerical, triangles), NLS estimator in (21) (WA03) (theoretical, dot-dashed line, and numerical, squares). The solid line black represents the CRB.*

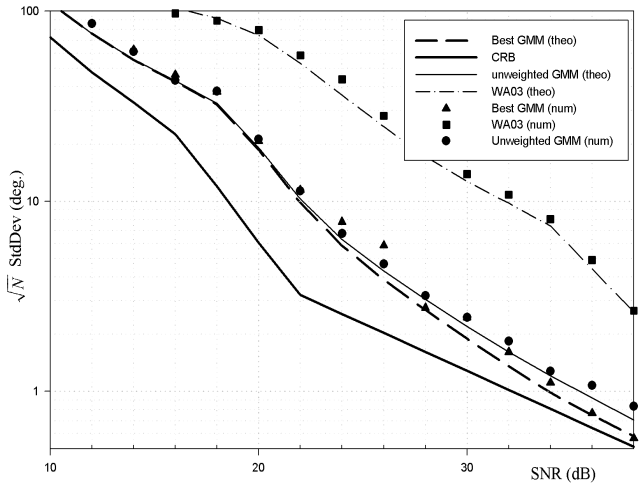

Figure 3.18: *Normalized root mean square phase estimation error* $\sqrt{N} \cdot RMSE$ *vs. SNR for 128-QAM constellation* $(N = 512, K = 2000)$*; unweighted GMM estimator (theoretical, solid line gray, and numerical, circles), optimal GMM estimator (theoretical, dashed line gray, and numerical, triangles), NLS estimator in (21) (WA03) (theoretical, dot-dashed line, and numerical, squares). The solid line black represents the CRB.*

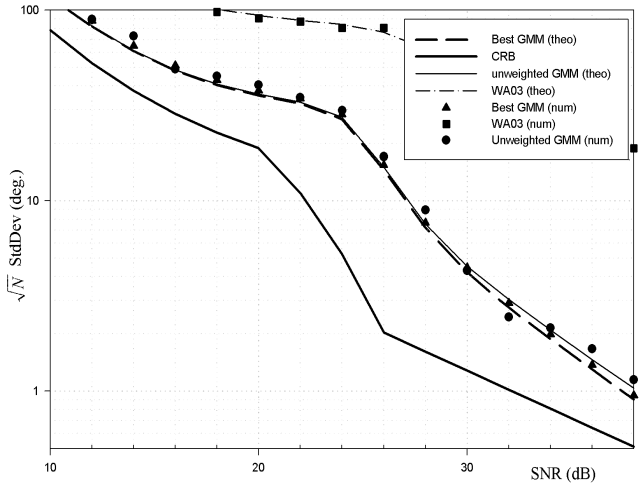

Figure 3.19: *Normalized root mean square phase estimation error* $\sqrt{N} \cdot RMSE$ *vs. SNR for 512-QAM constellation* $(N = 512, K = 2000)$*; unweighted GMM estimator (theoretical, solid line gray, and numerical, circles), optimal GMM estimator (theoretical, dashed line gray, and numerical, triangles), NLS estimator in (21) (WA03) (theoretical, dot-dashed line, and numerical, squares). The solid line black represents the CRB.*

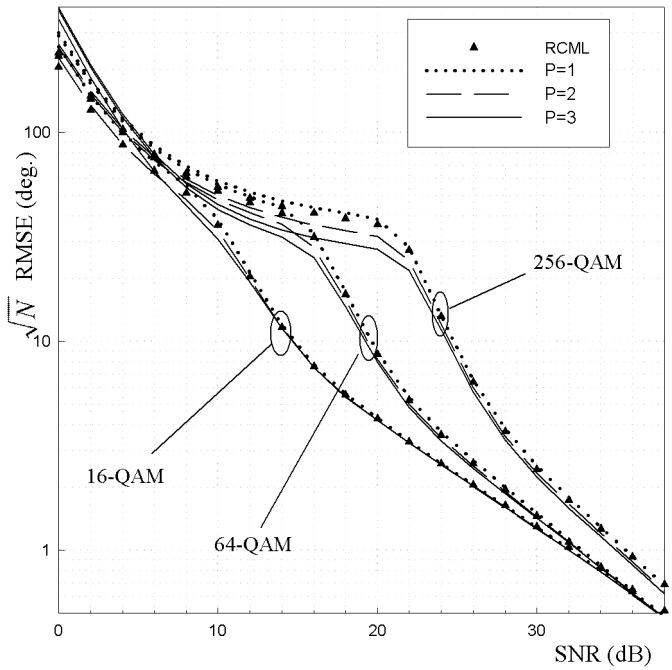

Figure 3.20: *Square constellations: analytical RMSE of the optimally weighted GMM phase estimator versus SNR for $P = 1$ (dotted line), $P = 2$ (dashed line), $P = 3$ (continuous line), and of the reduced complexity ML estimator (RCML) (triangle).*

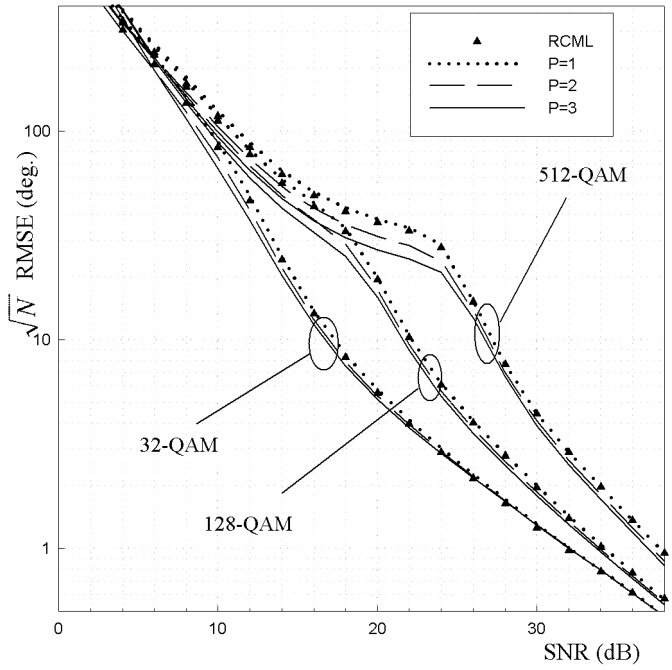

Figure 3.21: *Cross constellations: analytical RMSE of the optimally weighted GMM phase estimator versus SNR for $P = 1$ (dotted line), $P = 2$ (dashed line), $P = 3$ (continuous line), and of the reduced complexity ML estimator (RCML) (triangle).*

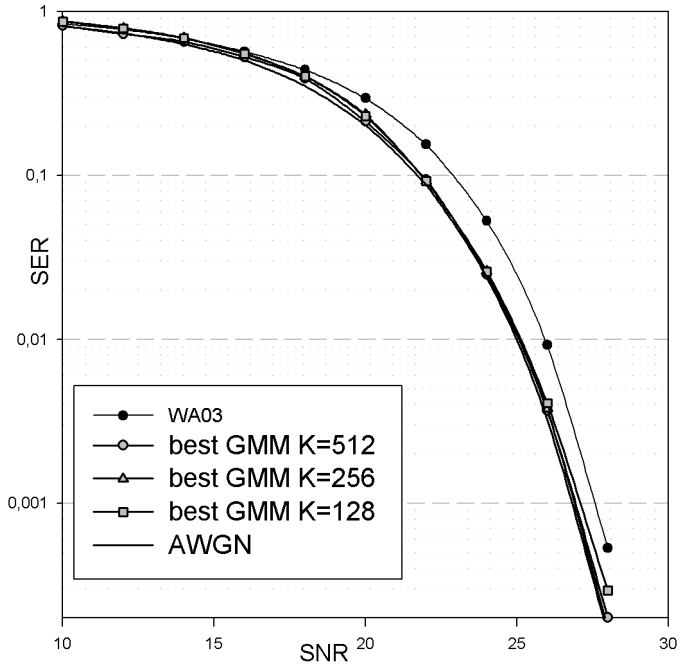

Figure 3.22: *SER vs. SNR for 128-QAM constellation* ($N = 512, K = 128, 256, 512$).

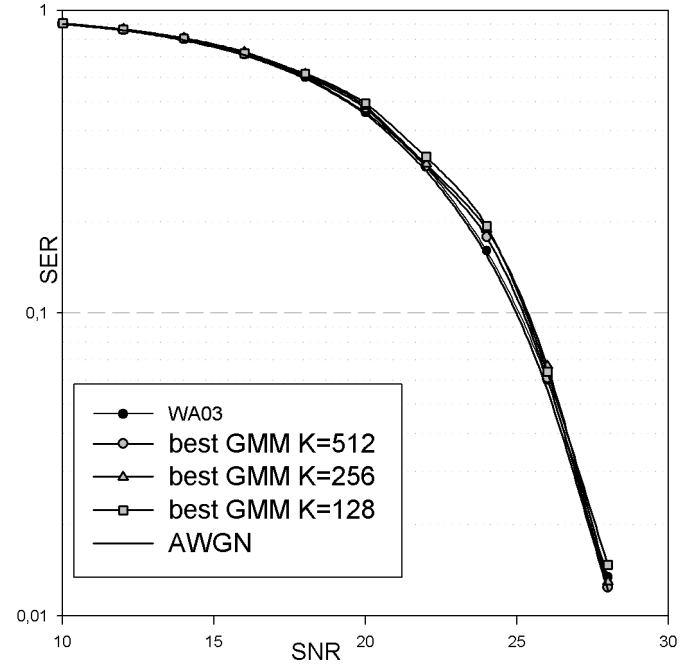

Figure 3.23: *SER vs. SNR for 256-QAM constellation* ($N = 512, K = 128, 256, 512$).

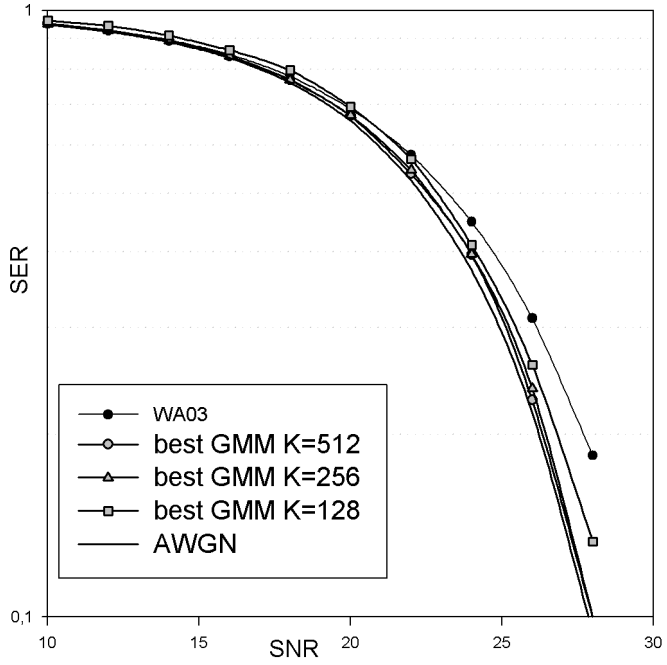

Figure 3.24: *SER vs. SNR for 512-QAM constellation* ($N = 512, K = 128, 256, 512$).

Chapter 4

GMM-based Frequency Offset Estimation: an Image Processing Approach

4.1 Introduction

In many Quadrature Amplitude Modulation transmission systems, Doppler effects and/or misalignments between local oscillators result in random carrier phase drifts. Hence, acquisition and tracking of carrier phase and frequency offset is typically performed at the output of the receiver front-end. The estimation of these parameters is carried out using training sequences, such as in many communication systems, or in a blind fashion, widely investigated to achieve bandwidth savings. In (37), a frequency estimator is derived by linearization of the Maximum-Likelihood (ML) cost function, both for preamble-based and blind acquisition. In (38), the authors analyze a frequency estimator based on the circumstance that the unique conjugate cyclic frequency of the received signal is equal to twice the frequency offset. Non-data aided frequency offset estimation is performed in (39) employing a suitable linear precoding, while in (40) the blind estimation exploits the unconjugated cyclostationarity of the fractionally sampled received signal statistics. In (21), the authors describe a joint carrier phase and frequency offset estimation method based on a Nonlinear Least-Squares (NLS) technique that aims at minimizing the asymptotic (large sample) error variance.

Such a NLS estimator, which is a generalized form of low Signal-to-Noise Ratio (SNR) ML estimator, requires the knowledge of the constellation and of both the channel gain and the SNR. In spite of this *a priori* information, its performance deteriorates for large constellation order and for cross-constellations. Hence, the problem of blind carrier parameters estimation for QAM transmission is still worth of investigations.

In this Chapter we extend the CPS-based phase offset estimator in Chapter 3, introducing a gain-control-free frequency offset estimator. Again the estimator procedure stems from the observation of peculiar structures of the constellation diagram in presence/absence of frequency offset, and employs image processing tools to exploit such regularities. Moreover, as for the phase offset introduced in the previous Chapter, it is possible to prove that the CPS-based estimation can be recast in the GMM framework.

Wherease the phase offset, produces a shift on the constellation diagram, because every point is subject to a constant phase rotation, the presence of a frequency drift induces a different phase shift, depending on the sampling time instant, for each constellation point, thus spreading all the points of the constellation diagram in the complex plane. Under these conditions, the phase histogram does not anymore exhibit the peaked shape of the CPS, the constellation points being no more clustered, but turns out to be a filtered version of the phase distribution corresponding to zero-frequency offset. Yet, if a frequency compensation is performed over the recevied samples, we expect that, the closer the compensation value gets to the real frequency offset, the more structured the constellation diagram will be after the compensation, all of the points being shifted back to their corresponding clusters. Under a perfect frequency compensation, the drift is completely removed and the constellation diagram exhibits the typical clusters of the reference zero frequency offset case; moreover the phase histogram tends again to the CPS. Hence, since the phase pdf corresponding to zero frequency offset is known for all the considered constellations, our novel frequency offset estimator considers the frequency compensation value that minimizes the distance between the measured weighed phase histogram and the weighed phase pdf calculated under the assumption of zero frequency offset. In this sense, we have recast the frequency offset estimation problem in terms of waveform matching between the *a priori* probabilistic description and a suitable *a posteriori* statistic.

It is worth noting that the objective function is gain independent and thus the frequency offset estimation procedure is gain-control-free.

We have also derived the analytical asymptotical (large number of samples) performance in high SNR, and assessed them by numerical simulations, showing that the herein described estimator outperforms state-of-the-art estimators for different constellations on a wide range of SNR values.

By the end of the Chapter we prove that the CPS-based frequency offset estimator can be seen as a GMM-based estimation technique.

4.2 Discrete-Time Signal Model

Let us consider a QAM transmission system, and let $S[n]$ be the n-th transmitted symbol drawn from a power normalized M-ary constellation $\mathcal{A}=\{S_0,\ldots,S_{M-1}\}$ with equiprobable symbols. At the receiver side, after front-end signal processing, a complex low-pass version of the received signal is sampled at symbol rate yielding the samples $X[n]$ for which we assume the following analytical model:

$$X[n] = G_C\, e^{j\theta+j2\pi f_0 n}\, S[n] + W[n] \tag{4.1}$$

where G_C is the gain, θ and f_0 are the phase and the frequency offset, and $W[n]$ is a realization of a circularly complex Gaussian stationary noise process, statistically independent of $S[n]$. The signal-to-noise ratio (SNR) is defined as $\mathrm{SNR} \stackrel{\mathrm{def}}{=} G_C^2/\sigma_W^2$, being $\sigma_W^2 \stackrel{\mathrm{def}}{=} \mathrm{E}\{|W[n]|^2\}$ the noise power.

In the following, we will address the estimation of the carrier frequency offset f_0 given a sample of N consecutive observations $X[n],\ n=0,\cdots,N-1$, with $|f_0|<1/8$,[1] and in absence of knowledge about the nuisance parameters G_C and θ.

[1]Due to the quadrant constellation symmetry, in absence of side information the phase of the signal samples cannot be recovered unless an ambiguity of $\pi/2$. A frequency offset $|f_0|>1/8$ corresponds to a phase jump larger than $\pm\pi/4$ for each received sample, that would be unrecoverable due to the said ambiguity.

4.3 Image Processing Approach to Phase-Offset Estimation

Let us now consider the following nonlinear function of the received signal samples $X[n]$ after a frequency compensation f_c:

$$Y^{(f_c)}[n] = |X[n]| \cdot e^{j4 \cdot \arg\{X[n]\}} e^{-j8\pi f_c n} \tag{4.2}$$

The nonlinearity (4.2) extends the transformation appearing in Section 3.3 and, for $f_c = f_0$, it folds the received samples whose phases differ by $\pi/2$. An analogous mapping is adopted in (37), where the received signal samples are projected in the first quadrant to take into account the constellation symmetry.

When the frequency compensation f_c perfectly removes the frequency offset, *i.e.* $f_c = f_0$, the nonlinearity (4.2) is exactly the same as in (3.2). In this case, all the observations in Sect.3.3 hold; we expect then the constellation diagram to exhibit a peculiar clusterization, depending on the constellation and on the SNR. Under this assumption, the MWTP of the bidimensional magnitude/phase pdf (3.5), exhibits a series of pulses, whose location and number depend on the constellation and whose shape depends on the SNR.

When instead no compensation is comprised, *i.e.* $f_c = 0$, each constellation point is shifted with respect to the preceeding because of the presence of the frequency offset f_0. Because of this, the points in the constellation diagram are spread over the complex plane, and the clusterization observed for perfect compensation vanishes. Consequently, the MWTP is expected to be a near uniform function, the pulses being spread over the phase axis as a consequence of the frequency offset.

In the noise free case, the pdf of the random variable $Y^{(f_c)}[n]$ is time-variant; adopting the polar representation $Y^{(f_c)}[n] = r[n]e^{j\varphi[n]}$, it assumes the following form involving a suite of Dirac pulses:

$$p_{R,\Phi}(r, \varphi; (f_c - f_0)n) = \frac{1}{M} \sum_{m=0}^{M-1} \delta(r - G_C|S_m|) \cdot \delta(\varphi - 4\theta + 8\pi(f_c - f_0)n - 4 \arg S_m) \tag{4.3}$$

where 4θ is the constant phase-shift due to the carrier phase-offset, and $8\pi(f_c - f_0)n$ is the time-variant phase-shift due to the residual carrier frequency offset $(f_c - f_0)$.

For perfect frequency compensation, *i.e.* $(f_c - f_0) = 0$, the noise-free pdf of $Y^{(f_c)}[n]$ becomes time-invariant and we re-obtain the expression in (3.4):

$$p_{R,\Phi}(r, \varphi; 0) = \frac{1}{M} \sum_{m=0}^{M-1} \delta(r - G_C|S_m|) \cdot \delta(\varphi - 4\theta - 4 \arg S_m) \tag{4.4}$$

In this case, the phase shift of the pdf of the random variable $Y^{(f_0)}[n]$ with respect to the variable φ reduces to the constant 4θ. Comparing (4.3) and (4.4), we observe that the following relation stands:

$$p_{R,\Phi}(r, \varphi; (f_c - f_0)n) = p_{R,\Phi}(r, \varphi - 8\pi(f_c - f_0)n; 0) \tag{4.5}$$

In presence of additive noise, the Dirac pulses appearing in (4.3) and (4.4) become wider pulses whose shape depends on the SNR and the noise pdf;[1] however, we can still observe the cyclic shift due to the residual frequency offset $(f_c - f_0)$, and the relation (4.5) still holds.

Fig.4.1 shows the nonlinearly transformed samples $Y^{(f_c)}[n]|_{f_c=0}$ observed in absence of frequency compensation as well as the nonlinearly transformed samples $Y^{(f_c)}[n]|_{f_c=f_0}$ observed in case of perfect frequency compensation (16-QAM, $N = 512, \mathrm{SNR} = 50\mathrm{dB}, f_0 = 0.005, \theta = 0$). For perfect frequency compensation, *i.e.* $(f_c - f_0) = 0$, the nonlinear transformation (4.2) folds signal samples originated by constellation points into a particular cluster of compensated samples $Y^{(f_c)}[n]|_{f_c=f_0}$.

Let us now suppose perfect frequency compensation $(f_c - f_0) = 0$ already accomplished, and let us consider the so-called Magnitude Weighed Tomographic Projection (MWTP) of the magnitude/phase bidimensional pdf $p_{R,\Phi}(r, \varphi; 0)$, as in (3.5):

$$g_{\Phi}^{(\mathcal{A},\theta,P)}(\varphi) \stackrel{\mathrm{def}}{=} \int_0^{+\infty} r \cdot p_{R,\Phi}(r, \varphi; 0)\, dr \tag{4.6}$$

[1] At high SNR, the shape of the pulses reduces to the pdf of the imaginary component of the noise $W[n]$.

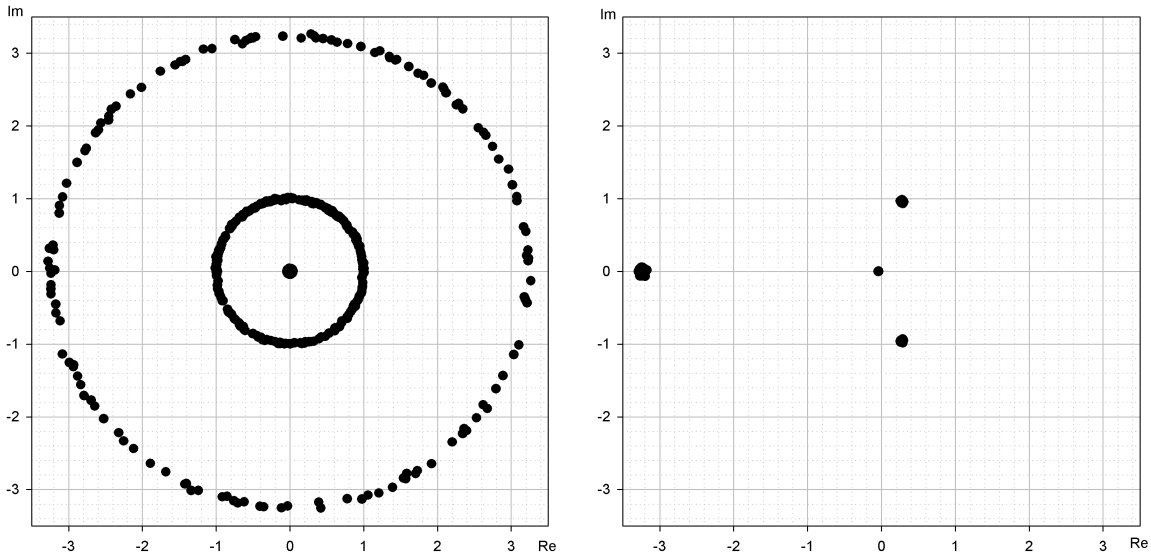

Figure 4.1: *Nonlinearly transformed samples* $Y^{(f_c)}[n]|_{f_c=0}$ *(no frequency compensation) and* $Y^{(f_c)}[n]|_{f_c=f_0}$ *(perfect frequency compensation), 16-QAM constellation (*$N = 512$*,* $SNR = 50$dB*,* $f_0 = 0.005$*,* $\theta = 0$*.)*

The expression of the zero phase-offset MWTP $g_{\Phi}^{(\mathcal{A},0,P)}(\varphi)$ is (cfr. App.E):

$$g_{\Phi}^{(\mathcal{A},0,P)}(\varphi) = \frac{4}{M\pi\,\sigma_W^2}\sum_{m=0}^{M-1}\exp\left(-\frac{G_C|S_m|^2\sin^2\left(\frac{\varphi}{4}-\arg S_m\right)}{\sigma_W^2}\right)$$
$$\cdot\sum_{k=0}^{2}\binom{2}{k}\left(G_C|S_m|\cos\left(\frac{\varphi}{4}-\arg S_m\right)\right)^{2-k}\int_{-G_C|S_m|\cos\left(\frac{\varphi}{4}-\arg S_m\right)}^{\infty}\rho^k\exp\left(-\frac{\rho^2}{\sigma_W^2}\right)d\rho \tag{4.7}$$

Fig.3.4 illustrates the MWTP for different QAM constellations (SNR = 35dB). It is worth noting that each MWTP shows a number of peaks equal to the number of clusters formed by perfectly compensated samples $Y^{(f_c)}[n]|_{f_c=f_0}$ originated by constellation points differing by $\pi/2$.

Following the same steps of Section 3.3 we evaluate the CPS as:

$$f^{(\mathcal{A},\theta,P)}(\psi) \stackrel{\text{def}}{=} \frac{K}{2\pi}\cdot\int_{\psi}^{\psi+2\pi/K} g_{\Phi}^{(\mathcal{A},\theta,P)}(\varphi)\;d\varphi \tag{4.8}$$

In the limit $K \to \infty$ it results:

$$\lim_{K\to\infty} f^{(\mathcal{A},\theta,P)}(\psi) = g_{\Phi}^{(\mathcal{A},\theta,P)}(\psi) \tag{4.9}$$

The CPS $f^{(\mathcal{A},\theta,P)}(\psi)$ of a QAM constellation $\mathcal{A}$ is formed by a set of pulses symmetrically located around $\pi + 4\theta$, whose relative distance and amplitudes depend on $\mathcal{A}$ and whose widths depend on SNR. In the CPS domain, as in MWTP domain, the phase offset θ is seen as a cyclic shift, *i.e.* $f^{(\mathcal{A},\theta,P)}(\psi) = f^{(\mathcal{A},0,P)}(\psi - 4\theta)$.

To illustrate the relationship between the nonlinearly transformed measurements $Y^{(f_c)}[n]$ and the CPS, let us consider K equiamplitude intervals

$$I_K^{(k)} \stackrel{\text{def}}{=} [2\pi k/K,\, 2\pi(k+1)/K)$$

for $k = 0, \ldots, K-1$, so as to introduce the following accumulation function:

$$a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k) \stackrel{\text{def}}{=} \frac{1}{N}\sum_{n=0}^{N-1} |Y^{(f_c)}[n]| \cdot d_K^{(k)}\left(Y^{(f_c)}[n]\right) \tag{4.10}$$

where $\psi_k \stackrel{\text{def}}{=} 2\pi k/K$ denotes the reference phase of the k-th phase interval and $d_K^{(k)}(Y)$ is an indicator function defined as follows:

$$d_K^{(k)}(Y) \stackrel{\text{def}}{=} \begin{cases} 1 & \arg Y \in I_K^{(k)} \\ 0 & \text{otherwise} \end{cases}$$

The accumulation function in (4.10) extends the sample CPS defined in (3.8). When perfect frequency compensation is accomplished, the accumulation function $a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)$ equals the form in (3.8). The expected value of $a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)$ is:

$$\begin{aligned}
\mathrm{E}\left\{a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)\right\} &= \frac{1}{N}\sum_{n=0}^{N-1} \mathrm{E}\left\{\left|Y^{(f_c)}[n]\right| \cdot d_K^{(k)}\left(Y^{(f_c)}[n]\right)\right\} \\
&= \frac{1}{N}\sum_{n=0}^{N-1} \int_0^{+\infty}\int_0^{2\pi} r\, p_{R,\Phi}(r,\varphi;(f_c-f_0)n)\, d_K^{(k)}\left(re^{j\varphi}\right)\, dr d\varphi \\
&= \frac{1}{N}\sum_{n=0}^{N-1} \int_{2\pi k/K}^{2\pi(k+1)/K}\int_0^{+\infty} r\, p_{R,\Phi}(r,\varphi;(f_c-f_0)n) dr_n\, d\varphi
\end{aligned}$$

Recalling that

$$p_{R,\Phi}(r,\varphi;(f_c-f_0)n) = p_{R,\Phi}(r,\varphi-8\pi(f_c-f_0)n;0)$$

we have:

$$\begin{aligned}
\mathrm{E}\left\{a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)\right\} &= \frac{1}{N}\sum_{n=0}^{N-1}\int_{2\pi k/K}^{2\pi(k+1)/K}\int_{0}^{+\infty} r\, p_{R,\Phi}(r,\varphi+8\pi(f_c-f_0)n;0)dr\, d\varphi \\
&= \frac{1}{N}\sum_{n=0}^{N-1}\int_{2\pi k/K+8\pi(f_c-f_0)n}^{2\pi(k+1)/K+8\pi(f_c-f_0)n}\int_{0}^{+\infty} r\, p_{R,\Phi}(r,\varphi;0)dr\, d\varphi \\
&= \frac{1}{N}\sum_{n=0}^{N-1} f^{(\mathcal{A},\theta,P)}(\psi_k+8\pi(f_c-f_0)n)
\end{aligned} \tag{4.11}$$

In (4.11) we recognize that, for $f_c-f_0\neq 0$, the expected value of the accumulation function $a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)$ is a superimposition of N suitably shifted versions of the CPS $f^{(\mathcal{A},\theta,P)}(\psi)$, *i.e.* $\mathrm{E}\left\{a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)\right\}$ is a filtered and sampled version of $f^{(\mathcal{A},\theta,P)}(\psi)$. On the other hand, in the case of perfect frequency compensation, $f_c-f_0=0$, (4.11) reduces to:

$$\mathrm{E}\left\{a^{(\mathcal{A},0,\theta)}(\psi_k)\right\} = f^{(\mathcal{A},\theta,P)}(\psi_k) \tag{4.12}$$

Hence, when perfect frequency compensation has been achieved, the accumulation function $a^{(\mathcal{A},0,\theta)}(\psi_k)$ turns out to be an unbiased estimate of the CPS $f^{(\mathcal{A},\theta,P)}(\psi_k)$.

Fig.4.2 gives a visual interpretation of these mathematical results about the accumulation function $a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)$, in the same case already addressed in Fig.4.1 (16-QAM constellation, $N=512$, SNR$=$50dB, $f_0=0.005$, $\theta=0$). The temporal evolution of the phase $\varphi[n]=\arg\left(Y^{(f_c)}[n]|_{f_c=0}\right)$ is represented in Fig.4.2(a) where a bidimensional phase-time scatter diagram is build by plotting the locus of points $(\varphi(n),n)$. In this phase-time domain, a cluster of the nonlinearly transformed samples $Y^{(f_c)}[n]$ appearing in Fig.4.1 corresponds to a linear pattern. In fact, due to the frequency offset f_0, the samples $Y^{(f_c)}[n]$ are affected by a linearly time-varying phase offset and, therefore, they result aligned along parallel linear patterns; the slope of the lines is proportional to the residual frequency offset f_c-f_0. For a perfect frequency offset compensation, the time-varying phase

offset vanishes and the linear patterns representing the phases of the samples $Y^{(f_c)}[n]|_{f_c=f_0}$ appear vertically oriented, as shown in Fig.4.2b.

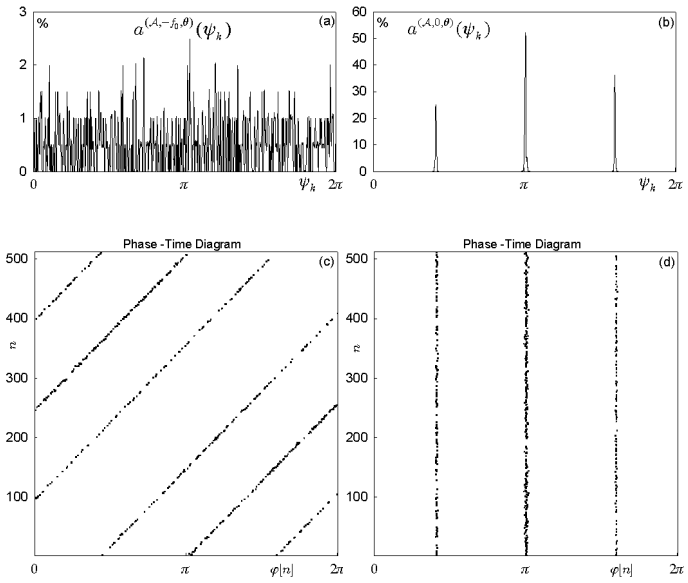

Figure 4.2: *Phase-time scatter plots of the nonlinearly transformed samples* $Y^{(f_c)}[n]$ *and corresponding sample functions* $a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)$ *(16-QAM constellation,* $N = 512$*,* $SNR = 50$dB*,* $f_0 = 0.005$*,* $\theta = 0$*). Subfigures (a) and (c) respectively show* $a^{(\mathcal{A},-f_0,\theta)}(\psi_k)$ *and* $Y^{(f_c)}[n]|_{f_c=0}$ *(no frequency compensation), subfigures (b) and (d) respectively show* $a^{(\mathcal{A},0,\theta)}(\psi_k)$ *and* $Y^{(f_c)}[n]|_{f_c=f_0}$ *(perfect frequency compensation).*

On the other hand, Figs.4.2c-d show the accumulation function $a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)$ evaluated in the cases of Figs.4.2a-b, respectively for $f_c-f_0=f_0$ and $f_c-f_0=0$.

In the first case, the samples corresponding to a specific cluster are spread out over different intervals $I_K^{(k)}$, and contribute to different samples of the accumulation function $a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)$, which appears uniform, although noisy, over the range $[0, 2\pi)$. Conversely, for perfect frequency compensation $f_c-f_0=0$, each cluster is represented by a vertical straight line and the accumulation function $a^{(\mathcal{A},0,\theta)}(\psi_k)$ is a peaked function that, for sufficiently large N, tends to the signal CPS.

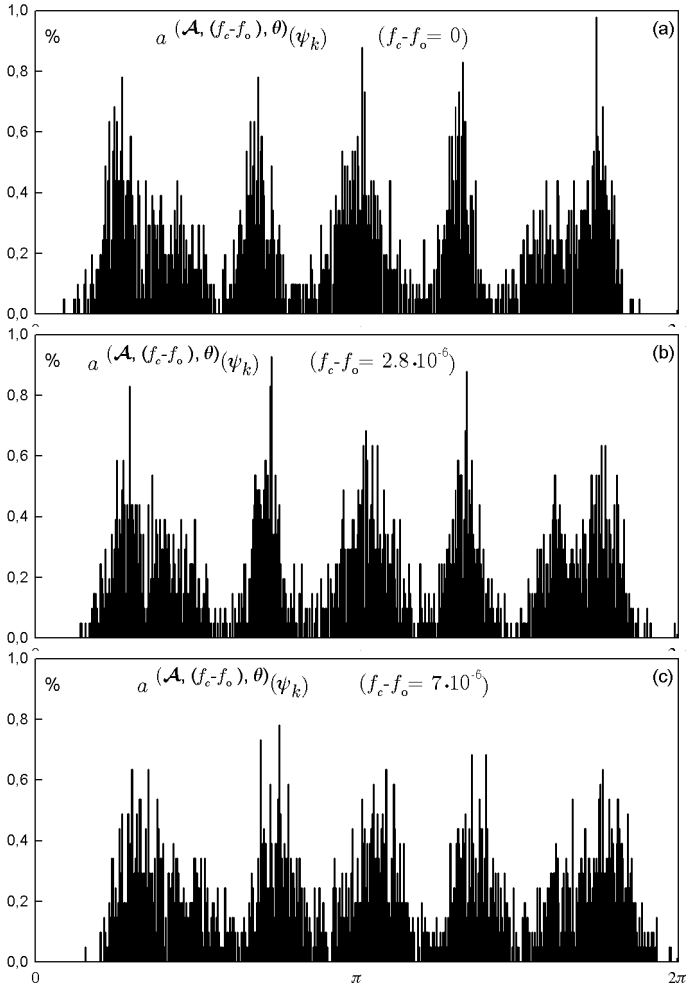

Figure 4.3: *Sample function* $a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)$ *evaluated at* $f_c-f_0=0$ *(a),* $f_c-f_0=2.8\cdot10^{-6}$ *(b), and* $f_c-f_0=7.0\cdot10^{-6}$ *(c) (32-QAM constellation SNR* $=23\text{dB}, \theta=0, N=2000, f_0=0.05$*).*

The deviation of the expected value of the accumulation function from the signal CPS is appreciated also for small values of the residual frequency offset f_c-f_0. Fig.4.3 plots the accumulation function $a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)$ for $f_c-f_0=0$, $f_c-f_0=2.8\cdot10^{-6}$, and $f_c-f_0=7.0\cdot10^{-6}$ in the case of 32-QAM ($N=2000$, $f_0=5\cdot10^{-2}$, SNR$=$23dB). Fig.4.3 shows that a small widening is discernible in $a^{(\mathcal{A},f_c-f_0,\theta)}(\psi_k)$ even in presence of relatively small values of the residual frequency offset f_c-f_0.

4.4 The CPS Based Frequency Offset Estimator

Bearing in mind the properties of the CPS above illustrated, it makes sense to estimate the frequency offset f_0 as the compensation value f_c that provides the best matching between the accumulation function $a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)$ and the reference CPS $f^{(\mathcal{A},\theta,P)}(\psi_k)$. Choosing a quadratic cost function, we may write:

$$\hat{f}_{CPS} = \arg\min_{f_c} \sum_{k=0}^{K-1} \left[a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k) - f^{(\mathcal{A},\theta,P)}(\psi_k)\right]^2 \tag{4.13}$$

In (4.13) the value of f_c affects both the energy of the measured accumulation function $a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)$

$$\mathcal{E}_a(f_c) \stackrel{\text{def}}{=} \sum_{k=0}^{K-1} a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)^2$$

and the cross-energy

$$\mathcal{C}(f_c) \stackrel{\text{def}}{=} \sum_{k=0}^{K-1} a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)\, f^{(\mathcal{A},\theta,P)}(\psi_k)\,.$$

Simulation results show that, within a relatively large interval around $f_c = f_0$, $\mathcal{E}_a(f_c)$ remains approximately constant[1]. Hence, the optimization in (4.13) is substantially obtained by maximizing the cross-energy term, which can be efficiently computed in the Discrete Fourier Transform (DFT) domain. Denoting by capital letter the DFT of a sequence, *e.g.* $A^{(\mathcal{A},(f_c-f_0),\theta)}(m) = ||\text{DFT}\{a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)\}||_m$, we can write:

$$\begin{aligned}\mathcal{C}(f_c) &= \sum_{m=0}^{K-1} A^{(\mathcal{A},(f_c-f_0),\theta)}(m) \cdot \overline{F^{(\mathcal{A},\theta)}(m)} \\ &= \sum_{m=0}^{K-1} A^{(\mathcal{A},(f_c-f_0),\theta)}(m) \cdot \left|F^{(\mathcal{A},0)}(m)\right| \cdot e^{j4\theta m}\end{aligned} \tag{4.14}$$

where the overline denotes complex conjugation. Since around f_0 the abscissa of the barycenter of $a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)$ well approximates θ, *cfr.* Section. 3.5, we can

[1]For instance, in the same cases addressed in Fig.4.3, the fluctuation of $\mathcal{E}_a(f_c)$ due to the different compensation values is less than the 1% of the fluctuation of $\mathcal{C}(f_c)$.

neglect the odd part of the sequence $a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)$ around its barycenter, thus obtaining:

$$\mathcal{C}(f_c) \simeq \sum_{m=0}^{K-1} \left|A^{(\mathcal{A},(f_c-f_0),\theta)}(m)\right| \cdot \left|F^{(\mathcal{A},0)}(m)\right| \tag{4.15}$$

Thus, we can conveniently adopt the following objective function:

$$\mathcal{G}(f_c) \stackrel{\text{def}}{=} \sum_{m=0}^{K-1} \left|A^{(\mathcal{A},(f_c-f_0),\theta)}(m)\right| \cdot \left|G_{\Phi}^{(\mathcal{A},0)}(m)\right| \tag{4.16}$$

where we have also fruitfully used the approximation $f^{(\mathcal{A},0,P)}(\psi_k) \simeq g_{\Phi}^{(\mathcal{A},0,P)}(\psi_k)$.

In summary, we have rephrased the frequency offset estimation stated in (4.13) as follows:

$$\hat{f}_{CPS} = \arg\max_{f_c} \{\mathcal{G}(f_c)\} \tag{4.17}$$

It is worthy noting that the estimator $\hat{f}_{CPS}$ does not need any preliminary gain control stage and/or phase offset compensation, but only a (coarse) SNR estimate for the CPS evaluation.[1]

The maximization of the objective function $\mathcal{G}(f_c)$ can be performed in a two-step coarse-to-fine fashion, *i.e.* by first evaluating an intermediate coarse estimate $\hat{f}_0^{(c)}$, and then by interpolating the function $\mathcal{G}(f_c)$ around $\hat{f}_0^{(c)}$ to obtain the fine estimate $\hat{f}_0^{(f)}$.

A two step approach is found, for instance, in (21). Therein, a coarse estimate of the frequency offset is first determined by applying a fast Fourier transform algorithm on the nonlinearly transformed, suitably zero-padded, measurements; thereafter, a fine frequency offset estimation is obtained by means of accurate interpolation in the frequency domain.

Here, the CPS-based coarse estimate $\hat{f}_0^{(c)}$is found by scanning the f_c admissible range $[-1/8, 1/8]$ with step Δf_c:

$$\begin{aligned} k_{max} &= \arg\max_{k} \{\mathcal{G}(k\Delta f_c)\} \\ \hat{f}_0^{(c)} &= k_{\max}\Delta f_c \end{aligned} \tag{4.18}$$

The computational complexity needed for scanning the entire admissible range of f_c can be further reduced by a preliminary scan based on a generic fast rough

[1]For a CPS based coarse SNR estimation see App. G.

estimator, thus limiting the CPS-based scanning into a small window around f_0. Even though any rough estimator could be used for this aim, for completeness, in App. I we provide a fast rough frequency estimator that still exploits the structure of the CPS; this rough estimator has been employed to limit the scanning range of the CPS-based coarse estimate $\hat{f}_0^{(c)}$ in the numerical experiments whose results are shown in Sect.4.6.

After the coarse estimation stage, we resort to a suitable interpolation techniques to refine the estimate $\hat{f}_0^{(c)}$, limited by the quantum value Δf_c. Following the approach presented in (15), we adopt a parabolic approximation for the objective function (4.16) around its maximum, thus obtaining the following expression for the fine estimate $\hat{f}_0^{(f)}$:

$$\hat{f}_0^{(f)} = \hat{f}_0^{(c)} - \frac{\Delta f_c}{8} \frac{\mathcal{G}(\hat{f}_0^{(c)} + \Delta f_c) - \mathcal{G}(\hat{f}_0^{(c)} - \Delta f_c)}{\mathcal{G}(\hat{f}_0^{(c)} + \Delta f_c) + \mathcal{G}(\hat{f}_0^{(c)} - \Delta f_c) - 2\mathcal{G}(\hat{f}_0^{(c)})} \tag{4.19}$$

This choice is motivated by the effectiveness of (4.19) in capturing the local variations of the function $\mathcal{G}(f_c)$ as well as by its analytical tractability, which allows to obtain the estimator performance in closed form. The following Section will be devoted to the asymptotical performance analysis of the frequency offset estimator (4.19).

4.5 High SNR Performance Analysis

In this Section we outline the high SNR performance analysis of the CPS based frequency offset estimator as given in (4.19). The estimation error is structured in two components: the first component occurs when the coarse estimate $\hat{f}_0^{(c)}$ is not correct, while the second error component is due to both the sample objective function estimation error and to parabolic approximation misfit around its maximum.

Numerical simulations show that, for medium to high SNR, the coarse estimate is always correct so that the estimator accuracy at high SNR is limited by the second error component only. Since the parabolic interpolation misfit can be reduced by using sufficiently small frequency interval Δf_c, the error on the fine estimate $\hat{f}_0^{(f)}$ is mainly due to the zero-mean sample estimation error of the objective function $\mathcal{G}(f_c)$; therefore, the bias of the fine estimate $\hat{f}_0^{(f)}$ due to the parabolic interpolation misfit is negligible and our theoretical analysis addresses the evaluation of the variance of $\hat{f}_0^{(f)}$.

Following the approach indicated in (15), the variance of $\hat{f}_0^{(f)}$ is analytically evaluated as a function of the mean, variance and covariances of the samples $\mathcal{G}(\hat{f}_0^{(c)})$, $\mathcal{G}(\hat{f}_0^{(c)} + \Delta f_c)$ and $\mathcal{G}(\hat{f}_0^{(c)} - \Delta f_c)$.

Let us set

$$x = \mathcal{G}(\hat{f}_0^{(c)} + \Delta f_c),\ y = \mathcal{G}(\hat{f}_0^{(c)} - \Delta f_c),\ z = \mathcal{G}(\hat{f}_0^{(c)})$$

and

$$X = \mathrm{E}\{x\},\ Y = \mathrm{E}\{y\},\ Z = \mathrm{E}\{z\}$$

Let us also compactly denote

$$c = X - Y,\ \ d = X - 2Z + Y$$

Then, within a first-order approximation of (4.19), the variance of the estimator $\hat{f}_0^{(f)}$ is given by:

$$\mathrm{Var}\{\hat{f}_0^{(f)}\} = \frac{\Delta f^2}{64}\Big[\left(\frac{d-c}{d^2}\right)^2 \mathrm{Var}\{x\} + \left(\frac{d+c}{d^2}\right)^2 \mathrm{Var}\{y\} + \left(\frac{2c}{d^2}\right)^2 \mathrm{Var}\{z\} - \left(\frac{d^2-c^2}{d^4}\right)\mathrm{Cov}\{x,y\} + \left(\frac{2dc+2c^2}{d^4}\right)\mathrm{Cov}\{z,y\} + \left(\frac{2dc-2c^2}{d^4}\right)\mathrm{Cov}\{x,z\}\Big] \tag{4.20}$$

From (4.20) we observe that the asymptotical frequency estimator accuracy is a function only of the first and second order moments of the sample objective function $\mathcal{G}(f_c)$. The evaluation of these moments is reported in App. H.

4.6 Numerical Experiments

Here, we assess the performance of the above described frequency offset estimation method by discussing numerical results obtained through numerical simulations. Each experiment consists of 1000 Monte Carlo trials using signal samples generated according to (4.1) with a sample size $N = 512$ for square-constellations (16-64-256 QAM) and $N = 2000$ for cross-constellations (32-128-512 QAM). Estimation is performed in two steps according to the coarse-to-fine approach described in Sect.4.3; the coarse frequency estimate is found as a multiple of the step Δf_c, using also the preliminary fast rough estimator described in App. I. For the sake of comparison, we have also reported the performances pertaining to the optimal Non-Linear-Square (NLS) estimator described in (21); to compare the here described CPS-based estimator with the NLS estimator, the frequency offset has been set at $f_0 = 0.05$.[1]

The results pertaining to the numerical experiments have been expressed in terms of the normalized Root Mean Square Error ($\sqrt{N} \cdot \mathrm{RMSE}$) measured on all the considered estimators; besides, since the bias of all the considered estimators is negligible or zero, the results pertaining to the theoretical asymptotic performance analysis have been rather expressed in terms of the normalized standard deviation of the estimation error $\sqrt{N} \cdot \mathrm{StdDev}\{\hat{f}_0^{(f)}\}$.

Figs. 4.4-4.9 show the analytical and experimental normalized RMSE versus the SNR. Both the CPS and the NLS estimators require the knowledge of the constellation and of the SNR. Simulation results have been obtained assuming perfect knowledge of this parameters at the receiver side. The NLS estimator requires also the knowledge of the gain whereas the CPS estimator is gain control free. The results pertaining to the NLS estimator have been obtained assuming also perfect gain knowledge. Furthermore, we have also reported the Cramér-Rao lower bound (CRB), as derived in Chapter 5.

[1] In (21), it has been also shown that, for $f_0 = 0.05$, the NLS estimator performs uniformly better than the estimator described in (28).

In Figs.4.4-4.9 we observe a good agreement between the high SNR asymptotical (large N) performance analysis in Sect. 4.5 and the numerical results either for square and cross-constellation, showing that the large sample approximation holds true for realistic values of N. At medium SNR the theoretical analysis slightly diverges from numerical results, still providing an upper bound to the numerical simulations performance.[1] At low SNR, the error of the coarse frequency estimate becomes dominant and the high SNR analysis is not applicable.

From Figs. 4.4-4.9 we see that for medium to high SNR ranges the CPS based estimator outperforms the NLS estimator and at high SNR it approaches the CRB. The performance achieved is explained by comparing the ML estimation and the CPS based estimation. At high SNR, after an ideal gain control stage, ML estimation minimizes the distance between every frequency compensated sample and the corresponding transmitted input symbols; the NLS estimator approximates the ML estimator for small order constellations. The CPS based estimation can be rather interpreted as the minimization of the difference between each peak of the accumulation function $a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k)$, rather representative of a whole *cluster* of received signal samples, and the peak of the CPS representing the subset of the corresponding transmitted symbols.

Since the CPS based estimator $\hat{f}_0^{(f)}$ requires the knowledge of SNR for the evaluation of the reference MWTP $g_{\Phi}^{(\mathcal{A},0,P)}(\psi_k)$, we complete the discussion presenting the effect of the SNR mismatch on the CPS based estimator accuracy. The performance degradation under the hypothesis of ± 5 dB and ± 10 dB SNR mismatch is shown in Figs.4.10 and 4.11 for 16-QAM and 32-QAM constellations, respectively. For each doublet of SNR mismatch, the curves report the worst measured performance. In both Figs.4.10 and 4.11, we can appreciate that the CPS based estimation accuracy preserves the CRB slope, although the mismatch slightly affects the estimator accuracy, resulting in a 2dB performance loss for a ± 5dB SNR mismatch up to 4dB for a ± 10dB SNR mismatch. The loss is restrained since the mismatch of the SNR parameter affects only the widths of the MWTP pulses. For the sake of comparison, we report also the performance of the optimal NLS estimator in (21) in case of SNR and/or gain mismatch. Tabs.4.1

[1] This circumstance arises because the variance in (4.20) has been derived without the constraint $\mathcal{G}(\hat{f}_0^{(c)}) > \mathcal{G}(\hat{f}_0^{(c)} \pm \Delta f_c)$, whereas this constraint is implicitly satisfied in the coarse-to-fine approach and it bounds the estimate $\hat{f}_0^{(f)}$ in the range $[\hat{f}_0^{(c)} - \Delta f_c/2, \hat{f}_0^{(c)} + \Delta f_c/2]$.

and 4.4 show the normalized RMSE of both the CPS and the NLS estimator for different values of SNR in presence of ±5dB and ±10dB SNR mismatch. Numerical results show that the CPS based estimator outperforms the optimal NLS estimator also in presence of SNR mismatch. Results reported in Tabs.4.2-4.6 comprise also a ±0.5dB gain mismatch. We observe that, especially in the case of cross-constellation, the presence of gain mismatch severely affects the performance of the NLS estimator in (21), while it has no effect on the CPS estimator.

For the sake of completeness, we also compared the herein introduced estimator with the gain-control-free sub-optimal fourth-order estimator in (21). Figs 4.12-4.13 show the analytical and experimental normalized RMSE for both the CPS and the sub-optimal fourth-order estimator in (21) along with the CRB for 16-QAM and 32-QAM constellations. Numerical results show that the CPS based estimator uniformly outperforms the fourth-order estimator in (21), especially in the case of cross-constellations. Moreover, the CPS based estimator does not show the saturation effect presented by the fourth-order estimator for $\mathrm{SNR} \to +\infty$.

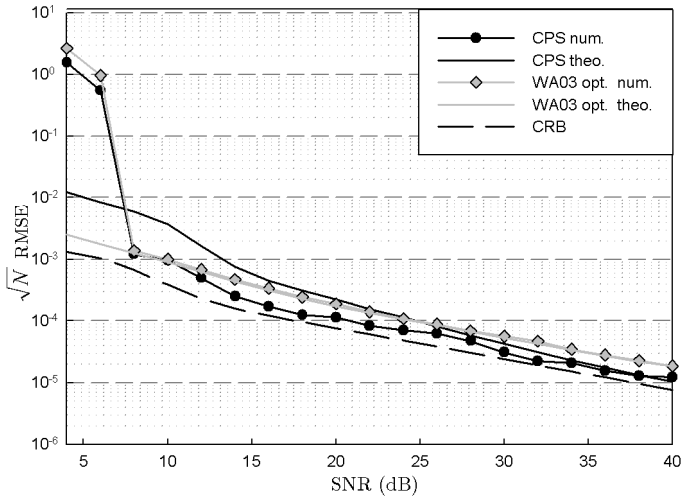

Figure 4.4: *Normalized root mean square frequency estimation error* $\sqrt{N} \cdot RMSE$ *vs. SNR for 16-QAM constellation* $(N = 512, K = 512)$*; CPS-based estimator (theoretical, solid line black, and numerical, circle black) and optimal NLS estimator in (21) (WA03) (theoretical, solid line gray, and numerical, diamond gray). The dashed line represents the CRB.*

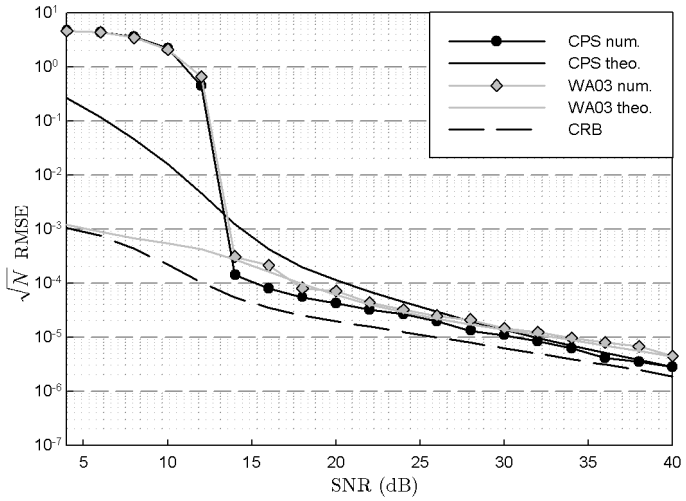

Figure 4.5: *Normalized root mean square frequency estimation error* $\sqrt{N} \cdot RMSE$ *vs. SNR for 32-QAM constellation* $(N = 2000, K = 512)$*; CPS-based estimator (theoretical, solid line black, and numerical, circle black) and optimal NLS estimator in (21) (WA03) (theoretical, solid line gray, and numerical, diamond gray). The dashed line represents the CRB.*

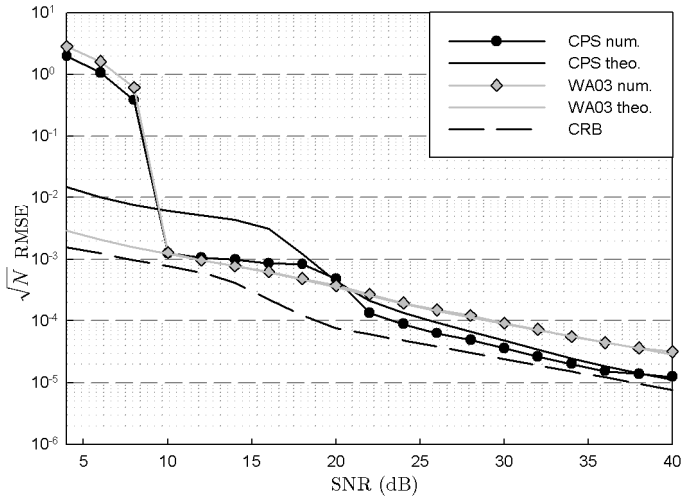

Figure 4.6: *Normalized root mean square frequency estimation error* $\sqrt{N} \cdot RMSE$ *vs. SNR for 64-QAM constellation* $(N = 512, K = 512)$*; CPS-based estimator (theoretical, solid line black, and numerical, circle black) and optimal NLS estimator in (21) (WA03) (theoretical, solid line gray, and numerical, diamond gray). The dashed line represents the CRB.*

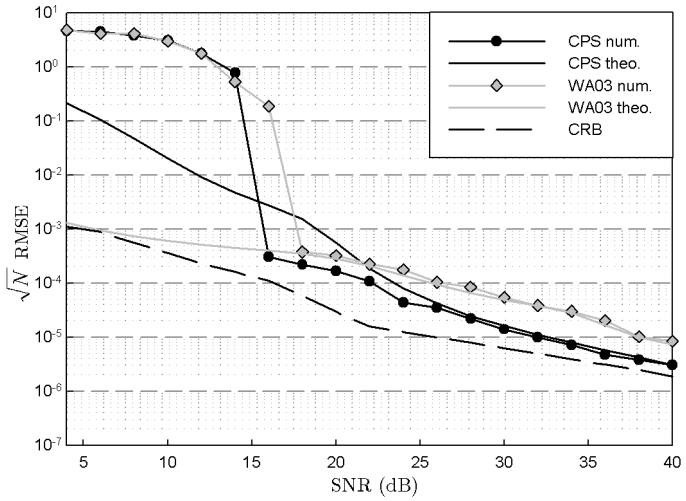

Figure 4.7: *Normalized root mean square frequency estimation error* $\sqrt{N} \cdot RMSE$ *vs. SNR for 128-QAM constellation* $(N = 2000, K = 512)$*; CPS-based estimator (theoretical, solid line black, and numerical, circle black) and optimal NLS estimator in (21) (WA03) (theoretical, solid line gray, and numerical, diamond gray). The dashed line represents the CRB.*

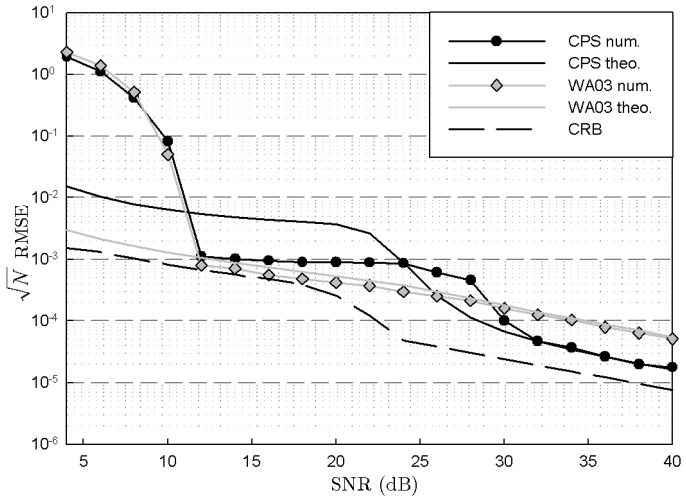

Figure 4.8: *Normalized root mean square frequency estimation error* $\sqrt{N} \cdot RMSE$ *vs. SNR for 256-QAM constellation* $(N = 512, K = 512)$*; CPS-based estimator (theoretical, solid line black, and numerical, circle black) and optimal NLS estimator in (21) (WA03) (theoretical, solid line gray, and numerical, diamond gray). The dashed line represents the CRB.*

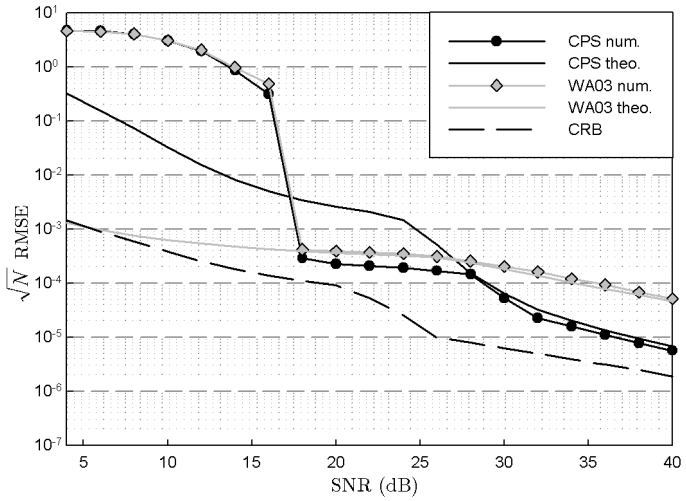

Figure 4.9: *Normalized root mean square frequency estimation error* $\sqrt{N} \cdot RMSE$ *vs. SNR for 512-QAM constellation* $(N = 2000, K = 512)$*; CPS-based estimator (theoretical, solid line black, and numerical, circle black) and optimal NLS estimator in (21) (WA03) (theoretical, solid line gray, and numerical, diamond gray). The dashed line represents the CRB.*

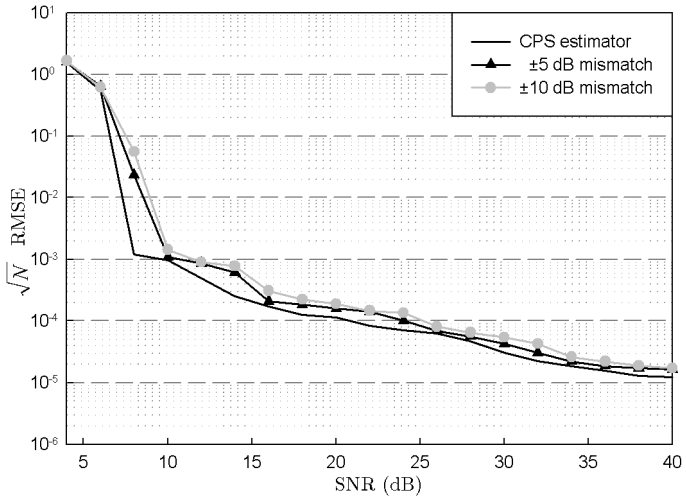

Figure 4.10: *Normalized root mean square frequency estimation error* $\sqrt{N} \cdot RMSE$ *vs. SNR for SNR mismatch equal to* ±5dB *(triangle) and* ±10dB *(circle) for 16 QAM constellation* ($N = 512, L = 512$). *For each doublet of SNR mismatch, the curves report the worst measured performance.*

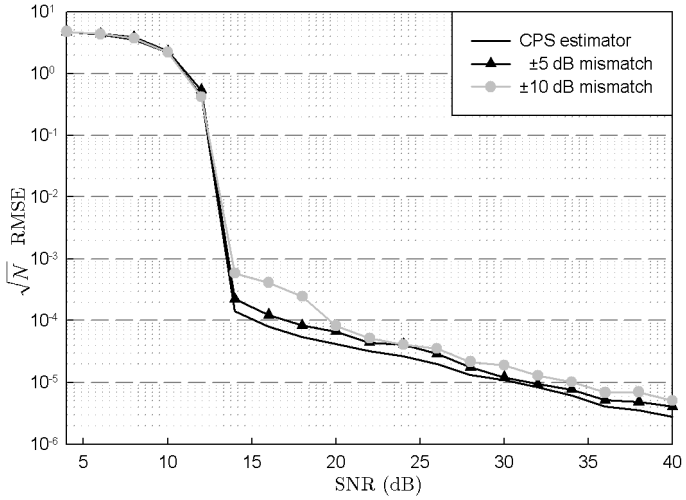

Figure 4.11: *Normalized root mean square frequency estimation error* $\sqrt{N} \cdot RMSE$ *vs. SNR for SNR mismatch equal to* ±5dB *(triangle) and* ±10dB *(circle) for 32 QAM constellation* ($N = 2000, L = 512$). *For each doublet of SNR mismatch, the curves report the worst measured performance.*

$\sqrt{N}$·RMSE 16QAM 0dB gain mism.	CPS estim.			WA03 estim.		
	SNR =20dB	SNR =30dB	SNR =40dB	SNR =20dB	SNR =30dB	SNR =40 dB
0dB SNR mismatch	$1.14 \cdot 10^{-4}$	$3.08 \cdot 10^{-5}$	$1.21 \cdot 10^{-5}$	$1.25 \cdot 10^{-4}$	$3.78 \cdot 10^{-5}$	$1.86 \cdot 10^{-5}$
±5dB SNR mismatch	$1.59 \cdot 10^{-4}$	$4.22 \cdot 10^{-5}$	$1.63 \cdot 10^{-5}$	$1.89 \cdot 10^{-4}$	$5.66 \cdot 10^{-5}$	$1.92 \cdot 10^{-5}$
±10dB SNR mismatch	$1.88 \cdot 10^{-4}$	$5.41 \cdot 10^{-5}$	$1.69 \cdot 10^{-5}$	$2.78 \cdot 10^{-4}$	$8.37 \cdot 10^{-5}$	$2.15 \cdot 10^{-5}$

Table 4.1:Normalized RMSE of CPS and WA03 estimators at different values of SNR, for ±5dB and ±10dB SNR mismatch, under a 0dB gain mismatch. (16-QAM constellation, $N = 512, K = 512$)

$\sqrt{N}$·RMSE 16QAM 0.5dB gain mism.	CPS estim.			WA03 estim.		
	SNR =20dB	SNR =30dB	SNR =40dB	SNR =20dB	SNR =30dB	SNR =40 dB
0dB SNR mismatch	$1.14 \cdot 10^{-4}$	$3.08 \cdot 10^{-5}$	$1.21 \cdot 10^{-5}$	$2.06 \cdot 10^{-4}$	$6.45 \cdot 10^{-5}$	$2.20 \cdot 10^{-5}$
±5dB SNR mismatch	$1.59 \cdot 10^{-4}$	$4.22 \cdot 10^{-5}$	$1.63 \cdot 10^{-5}$	$3.05 \cdot 10^{-4}$	$7.81 \cdot 10^{-5}$	$2.30 \cdot 10^{-5}$
±10dB SNR mismatch	$1.88 \cdot 10^{-4}$	$5.41 \cdot 10^{-5}$	$1.69 \cdot 10^{-5}$	$3.36 \cdot 10^{-4}$	$9.55 \cdot 10^{-5}$	$3.11 \cdot 10^{-5}$

Table 4.2: Normalized RMSE of CPS and WA03 estimators at different values of SNR, for ±5dB and ±10dB SNR mismatch, under a 0.5dB gain mismatch. (16-QAM constellation, $N = 512, K = 512$)

$\sqrt{N}$·RMSE 16QAM -0.5dB gain mism.	CPS estim.			WA03 estim.		
	SNR =20dB	SNR =30dB	SNR =40dB	SNR =20dB	SNR =30dB	SNR =40 dB
0dB SNR mismatch	$1.14 \cdot 10^{-4}$	$3.08 \cdot 10^{-5}$	$1.21 \cdot 10^{-5}$	$1.64 \cdot 10^{-4}$	$5.17 \cdot 10^{-5}$	$2.26 \cdot 10^{-5}$
±5dB SNR mismatch	$1.59 \cdot 10^{-4}$	$4.22 \cdot 10^{-5}$	$1.63 \cdot 10^{-5}$	$2.01 \cdot 10^{-4}$	$7.76 \cdot 10^{-5}$	$2.42 \cdot 10^{-5}$
±10dB SNR mismatch	$1.88 \cdot 10^{-4}$	$5.41 \cdot 10^{-5}$	$1.69 \cdot 10^{-5}$	$5.52 \cdot 10^{-4}$	$9.22 \cdot 10^{-5}$	$3.71 \cdot 10^{-5}$

Table 4.3: Normalized RMSE of CPS and WA03 estimators at different values of SNR, for ±5dB and ±10dB SNR mismatch, under a -0.5dB gain mismatch. (16-QAM constellation, $N = 512, K = 512$)

$\sqrt{N}\cdot$RMSE 32QAM 0dB gain mism.	CPS estim.			WA03 estim.		
	SNR =20dB	SNR =30dB	SNR =40dB	SNR =20dB	SNR =30dB	SNR =40 dB
0dB SNR mismatch	$4.19 \cdot 10^{-5}$	$1.80 \cdot 10^{-5}$	$2.79 \cdot 10^{-6}$	$7.13 \cdot 10^{-5}$	$1.43 \cdot 10^{-5}$	$4.46 \cdot 10^{-6}$
±5dB SNR mismatch	$6.66 \cdot 10^{-5}$	$1.20 \cdot 10^{-5}$	$4.03 \cdot 10^{-6}$	$7.99 \cdot 10^{-4}$	$7.94 \cdot 10^{-4}$	$8.00 \cdot 10^{-4}$
±10dB SNR mismatch	$8.23 \cdot 10^{-5}$	$1.88 \cdot 10^{-5}$	$5.00 \cdot 10^{-6}$	$7.95 \cdot 10^{-4}$	$8.06 \cdot 10^{-4}$	$8.04 \cdot 10^{-4}$

Table 4.4:Normalized RMSE of CPS and WA03 estimators at different values of SNR, for ±5dB and ±10dB SNR mismatch, under a 0dB gain mismatch. (32-QAM constellation, $N = 2000, K = 512$)

$\sqrt{N}\cdot$RMSE 32QAM 0.5dB gain mism.	CPS estim.			WA03 estim.		
	SNR =20dB	SNR =30dB	SNR =40dB	SNR =20dB	SNR =30dB	SNR =40 dB
0dB SNR mismatch	$4.19 \cdot 10^{-5}$	$1.80 \cdot 10^{-5}$	$2.79 \cdot 10^{-6}$	$1.59 \cdot 10^{-3}$	$1.58 \cdot 10^{-3}$	$1.55 \cdot 10^{-3}$
±5dB SNR mismatch	$6.66 \cdot 10^{-5}$	$1.20 \cdot 10^{-5}$	$4.03 \cdot 10^{-6}$	1.56	$1.57 \cdot 10^{-3}$	$1.59 \cdot 10^{-3}$
±10dB SNR mismatch	$8.23 \cdot 10^{-5}$	$1.88 \cdot 10^{-5}$	$5.00 \cdot 10^{-6}$	1.43	$1.56 \cdot 10^{-3}$	$1.59 \cdot 10^{-3}$

Table 4.5: Normalized RMSE of CPS and WA03 estimators at different values of SNR, for ±5dB and ±10dB SNR mismatch, under a 0.5dB gain mismatch. (32-QAM constellation, $N = 2000, K = 512$)

$\sqrt{N}\cdot$RMSE 32QAM -0.5dB gain mism.	CPS estim.			WA03 estim.		
	SNR =20dB	SNR =30dB	SNR =40dB	SNR =20dB	SNR =30dB	SNR =40 dB
0dB SNR mismatch	$4.19 \cdot 10^{-5}$	$1.80 \cdot 10^{-5}$	$2.79 \cdot 10^{-6}$	4.67	1.51	1.41
±5dB SNR mismatch	$6.66 \cdot 10^{-5}$	$1.20 \cdot 10^{-5}$	$4.03 \cdot 10^{-6}$	4.47	3.99	1.29
±10dB SNR mismatch	$8.23 \cdot 10^{-5}$	$1.88 \cdot 10^{-5}$	$5.00 \cdot 10^{-6}$	4.48	1.53	1.62

Table 4.6: Normalized RMSE of CPS and WA03 estimators at different values of SNR, for ±5dB and ±10dB SNR mismatch, under a -0.5dB gain mismatch. (32-QAM constellation, $N = 2000, K = 512$)

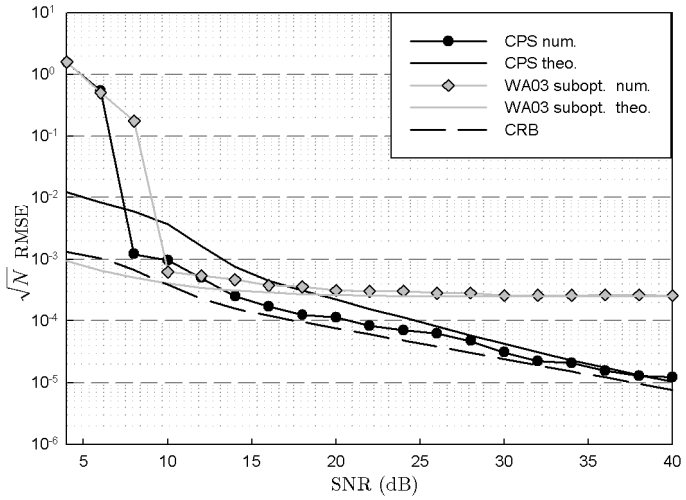

Figure 4.12: *Normalized root mean square frequency estimation error* $\sqrt{N} \cdot RMSE$ *vs. SNR for 16-QAM constellation* ($N = 512, K = 512$)*; CPS-based estimator (theoretical, solid line black, and numerical, circle black) and sub-optimal fourth-order estimator in (21) (WA03) (theoretical, solid line gray, and numerical, diamond gray). The dashed line represents the CRB.*

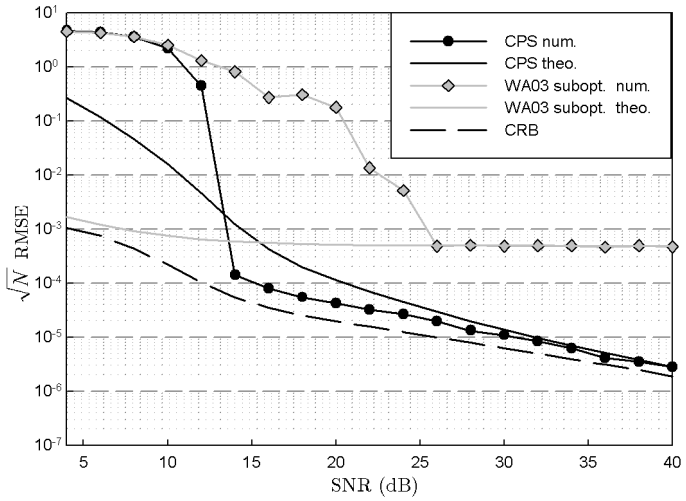

Figure 4.13: *Normalized root mean square frequency estimation error* $\sqrt{N} \cdot RMSE$ *vs. SNR for 32-QAM constellation* ($N = 2000, K = 512$)*; CPS-based estimator (theoretical, solid line black, and numerical, circle black) and sub-optimal fourth-order estimator in (21) (WA03) (theoretical, solid line gray, and numerical, diamond gray). The dashed line represents the CRB.*

4.7 GMM based Frequency Acquisition for QAM Constellation

In this Section we briefly outline how also the CPS-based frequency offset acquisition introduced in Sect. 4.4, can be recst in the framework of the GMM.

Let us in fact consider the estimation rule as in (4.13):

$$\hat{f}_{CPS} = \arg\min_{f_c} \sum_{k=0}^{K-1} \left[a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_k) - f^{(\mathcal{A},\theta,P)}(\psi_k) \right]^2 \tag{4.21}$$

By defining the following observations and reference vectors as in (3.20) and (3.21)

$$\mathbf{a}(f_c,\theta) = \left[a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_0), \ldots, a^{(\mathcal{A},(f_c-f_0),\theta)}(\psi_{K-1}) \right]^{\mathrm{T}}$$

$$\mathbf{f}(\xi) = \left[f^{(\mathcal{A},\xi,P)}(\psi_0), \ldots, f^{(\mathcal{A},\xi,P)}(\psi_{K-1}) \right]^{\mathrm{T}}$$

and by considering the following generalized moment:

$$\mathbf{e}(f_c) \stackrel{\mathrm{def}}{=} \mathbf{a}(f_c,\theta) - \mathbf{f}(0)$$

we can re-write the estimator in (4.21) as the solution of the following unweighted GMM estimation problem:

$$\hat{f}_{CPS} = \arg\min_{f_c} \mathbf{e}(f_c)^{\mathrm{T}} \mathbf{e}(f_c) \tag{4.22}$$

The time varying nature of the shift induced by the frequency offset in (4.3) makes unfeasible the estimation approach for location parameters introduced in Chapter 2. More specifically the extension of 4.22 to the optimally weighted case presents serious computational issues, due to the inversion of the measurements covariance matrix for every estimation step.

The analysis of a domain so to recast the frequecny offset estimation as a location parameter is hence left for future study [1].

[1] An example in this sense may be given by the Radon transformed domain

Chapter 5

Cramér Rao Lower Bound for QAM Phase and Frequency Offset Estimation

5.1 Introduction

In this Chapter we detail the derivation of the Cramér Rao Lower Bound (CRLB) for the estimation of phase and frequency offset for general QAM constellation, under the hypothesis that the gain factor is unknown, and hence to be estimated itself. To evaluate the score functions we follow the approach in (12), and to perform the expectations for the evaluation of the Fisher information matrix we resort to a Monte Carlo integration technique (49).

5.2 Cramér Rao Lower Bound for joint Phase and Gain Estimation

Let us consider the same discrete-time signal model employed in Chpater 3, that is:

$$X_n = G_C\, e^{j\theta}\, S_n + W_n \tag{5.1}$$

We recall that G_C is the unknown overall gain, θ is the unknown phase-offset, and W_n is a realization of a circularly complex Gaussian stationary noise pro-

cess, statistically independent of S_n. The signal-to-noise ratio (SNR) is defined as $\mathrm{SNR} \stackrel{\mathrm{def}}{=} G_C^2/\sigma_W^2$, being $\sigma_W^2 \stackrel{\mathrm{def}}{=} \mathrm{E}\{|W_n|^2\}$ the noise variance. The S_n are the transmitted symbols drawn from a power normalized QAM constellation $\mathcal{A}$. For instance, in the reference case of the 16 QAM constellation we have that:

$$\mathcal{A}_{16} = \frac{1}{\sqrt{10}}\{\pm 1 \pm j, \pm 1 \pm 3j, \pm 3 \pm j, \pm 3 \pm 3j\}$$

Th Cramér Rao Lower Bound (CRB) for the variance of an unbiased estimator $\hat{\theta}$ is well known to be given by (47):

$$\mathrm{CRB}\left(\hat{\theta}\right) = -\left[\mathrm{E}\left\{\frac{\partial^2 \ln p\left(\mathbf{X}|\theta\right)}{\partial \theta^2}\right\}\right]^{-1}$$

where $p\left(\mathbf{X}|\theta\right)$ denotes the conditional pdf of the observed sample $\mathbf{X} = \{X_i\}_{i=0}^{N-1}$. When considering the case of joint phase and gain estimation, the expression the CRB for the phase offset θ estimation is found as the element $(1,1)$ of the inverse Fisher Information Matrix as defined as follows:

$$\mathrm{CRB}\left(\hat{\theta}\right) = \left[\mathbf{F}\left(\theta,\ G_C\right)^{-1}\right]_{(1,1)} \tag{5.2}$$

being:

$$\mathbf{F}\left(\theta,\ G_C\right) = \begin{bmatrix} -\mathrm{E}\left\{\frac{\partial^2 \ln p(\mathbf{X}|\theta,\ G_C)}{\partial \theta^2}\right\} & -\mathrm{E}\left\{\frac{\partial^2 \ln p(\mathbf{X}|\theta,\ G_C)}{\partial \theta\ \partial G_C}\right\} \\ & \\ -\mathrm{E}\left\{\frac{\partial^2 \ln p(\mathbf{X}|\theta,\ G_C)}{\partial \theta\ \partial G_C}\right\} & -\mathrm{E}\left\{\frac{\partial^2 \ln p(\mathbf{X}|\theta,\ G_C)}{\partial G_C^2}\right\} \end{bmatrix} \tag{5.3}$$

For a single received sample X_0, the corresponding conditional pdf $p\left(X_0|\theta,\ G_C\right)$ is given by:

$$p\left(X_0|\theta,\ G_C\right) = \sum_{S_i \in \mathcal{A}} \frac{P\left(S_i\right)}{2\pi\sigma_W^2} e^{-\frac{|X_0 - G_C S_i e^{j\theta}|^2}{2\sigma_W^2}}$$

being $P\left(S_i\right)$ the probability of emission of the constellation symbol S_i.[1] Since the recevied samples are statistycally independent, the joint pdf of the N recevied samples in $\mathbf{X}$ is given by:

$$p\left(\mathbf{X}|\theta,\ G_C\right) = \prod_{n=0}^{N-1} \left\{ \sum_{S_i \in \mathcal{A}} \frac{1}{2\pi M\sigma_W^2} e^{-\frac{\left|X_n - G_C S_i e^{j\theta}\right|^2}{2\sigma_W^2}} \right\} \tag{5.4}$$

[1] From now on, without loss of generality we will consider the constellation symbols to be equiprobable so that, for a generic M-QAM constellation we have that $P(S_i) = \frac{1}{M}$.

Expanding (5.4), and following the same approach as in (12), we come up with the following steps:

$$p\left(\mathbf{X}|\theta,\ G_C\right) = \prod_{n=0}^{N-1}\left\{\frac{1}{2\pi M\sigma_W^2}e^{-\frac{|X_n|^2}{2\sigma_W^2}}\sum_{S_i\in\mathcal{A}}e^{-\frac{|G_CS_i|^2}{2\sigma_W^2}}e^{\frac{\Re\left[X_nG_CS_i^*e^{-j\theta}\right]}{\sigma_W^2}}\right\}$$

$$= \prod_{n=0}^{N-1}\left\{\frac{1}{2\pi M\sigma_W^2}e^{-\frac{|X_n|^2}{2\sigma_W^2}}\sum_{S_i\in\mathcal{Q}_1}e^{-\frac{|G_CS_i|^2}{2\sigma_W^2}}\cdot\sum_{r=\pm1}\cosh\left[\Re\left(\frac{1}{\sigma_W^2}\chi\left(\theta,\ G_C,\ n,\ i,\ r\right)\right)\right]\right\} \tag{5.5}$$

being:

$$\chi\left(\theta,\ G_C,\ n,\ i,\ r\right)\stackrel{\text{def}}{=}|G_CS_i|\,X_ne^{-j\left(\theta+r\arctan\left(\frac{\Im\left(S_i\right)}{\Re\left(S_i\right)}\right)\right)} \tag{5.6}$$

where the superscript $(\cdot)^*$ stands for the complex conjugate, $\Re(\cdot)$ and $\Im(\cdot)$ for real and imaginary parts, and $\mathcal{Q}_1$ consists of the cosntellation points in the first quadrant. The derivation of (5.14) can be found in the Appendix of (12).

Stemming from the definition of $\chi\left(\theta,\ G_C,\ n,\ i,\ r\right)$ in (5.6), we have that the following relations stand:

$$\frac{\partial}{\partial\theta}\Re\left[\chi\left(\theta,\ G_C,\ n,\ i,\ r\right)\right] = \Im\left[\chi\left(\theta,\ G_C,\ n,\ i,\ r\right)\right]$$

$$\frac{\partial}{\partial\theta}\Im\left[\chi\left(\theta,\ G_C,\ n,\ i,\ r\right)\right] = -\Re\left[\chi\left(\theta,\ G_C,\ n,\ i,\ r\right)\right]$$

$$\frac{\partial}{\partial G_C}\Re\left[\chi\left(\theta,\ G_C,\ n,\ i,\ r\right)\right] = \Re\left[\frac{\chi\left(\theta,\ G_C,\ n,\ i,\ r\right)}{G_C}\right] = \Re\left[\chi_\theta\left(\theta,\ n,\ i,\ r\right)\right]$$

$$\frac{\partial}{\partial G_C}\Im\left[\chi\left(\theta,\ G_C,\ n,\ i,\ r\right)\right] = \Im\left[\frac{\chi\left(\theta,\ G_C,\ n,\ i,\ r\right)}{G_C}\right] = \Im\left[\chi_\theta\left(\theta,\ n,\ i,\ r\right)\right]$$

Taking the logarithm and differentiating twice (5.14), we obtain the following expressions for the second partial derivatives in (5.11):

$$\frac{\partial^2}{\partial\theta^2}\ln p\left(\mathbf{X}|\theta,\ G_C\right)=\sum_{n=0}^{N-1}\Bigg\{$$
$$\left[\sum_{S_i\in\mathcal{Q}_1}e^{-\frac{\alpha^2}{2}}\left(\frac{1}{\sigma_W^4}\sum_{r=\pm1}\cosh\left(R\right)\Im^2\left(\chi\right)-\frac{1}{\sigma_W^2}\sum_{r=\pm1}\sinh\left(R\right)\Re\left(\chi\right)\right)\right]\left[D\left(\theta,\ G_C,\ X_n,\ \mathcal{Q}_i,\sigma_W^2\right)\right]^{-1}$$
$$-\left[\sum_{S_i\in\mathcal{Q}_1}e^{-\frac{\alpha^2}{2}}\left(\frac{1}{\sigma_W^2}\sum_{r=\pm1}\sinh\left(R\right)\Im\left(\chi\right)\right)\right]^2\left[D\left(\theta,\ G_C,\ X_n,\ \mathcal{Q}_i,\sigma_W^2\right)\right]^{-2}\Bigg\} \tag{5.7}$$

$$\frac{\partial^2}{\partial\theta\ \partial G_C}\ln p\left(\mathbf{X}|\theta,\ G_C\right)=$$
$$\sum_{n=0}^{N-1}\Bigg\{\Bigg[\sum_{S_i\in\mathcal{Q}_1}e^{-\frac{\alpha^2}{2}}\Bigg(-\frac{G_C\left|S_i\right|^4}{\sigma_W^4}\sum_{r=\pm1}\sinh\left(R\right)\Re\left(\chi\left(\theta,\ G_C,\ n,\ i,\ r\right)\right)$$
$$+\frac{1}{\sigma_W^4}\sum_{r=\pm1}\cosh\left(R\right)\Re\left(\chi_\theta\left(\theta,\ n,\ i,\ r\right)\right)\Im\left(\chi\left(\theta,\ G_C,\ n,\ i,\ r\right)\right)$$
$$+\frac{1}{\sigma_W^2}\sum_{r=\pm1}\sinh\left(R\right)\Im\left(\chi_\theta\left(\theta,\ n,\ i,\ r\right)\right)\Bigg)\Bigg]\left[D\left(\theta,\ G_C,\ X_n,\ \mathcal{Q}_i,\sigma_W^2\right)\right]^{-1}$$
$$-\left[\sum_{S_i\in\mathcal{Q}_1}\frac{e^{-\frac{\alpha^2}{2}}}{\sigma_W^2}\sum_{r=\pm1}\sinh\left(R\right)\Im\left(\chi\left(\theta,\ G_C,\ n,\ i,\ r\right)\right)\right]\cdot\left[\sum_{S_i\in\mathcal{Q}_1}\frac{e^{-\frac{\alpha^2}{2}}}{\sigma_W^2}\right.$$
$$\left.\left(-G_C\left|S_i\right|\sum_{r=\pm1}\cosh\left(R\right)+\sum_{r=\pm1}\sinh\left(R\right)\Re\left(\chi_\theta\left(\theta,\ n,\ i,\ r\right)\right)\right)\right]\left[D\left(\theta,\ G_C,\ X_n,\ \mathcal{Q}_i,\sigma_W^2\right)\right]^{-2}\Bigg\} \tag{5.8}$$

and

$$\frac{\partial^2}{\partial G_C^2}\ln p\left(\mathbf{X}|\theta,\ G_C\right) =$$
$$\sum_{n=0}^{N-1}\left\{\left[\sum_{S_i\in\Omega_1}\frac{e^{-\frac{\alpha^2}{2}}}{\sigma_W^2}\left(|S_i|^2\left(\alpha^2-1\right)\sum_{r=\pm1}\cosh\left(R\right)+\frac{1}{\sigma_W^2}\sum_{r=\pm1}\cosh\left(R\right)\Re\left(\chi_\theta\left(\theta,\ n,\ i,\ r\right)\right)^2\right.\right.$$
$$\left.\left.-G_C\left|S_i\right|^2\left(1+\frac{1}{\sigma_W^2}\right)\sum_{r=\pm1}\sinh\left(R\right)\Re\left(\chi_\theta\left(\theta,\ n,\ i,\ r\right)\right)\right)\right]\cdot\left[D\left(\theta,\ G_C,\ X_n,\ \Omega_i,\sigma_W^2\right)\right]^{-1}$$
$$-\left[\sum_{S_i\in\Omega_1}\frac{e^{-\frac{\alpha^2}{2}}}{\sigma_W^2}\left(-G_C\left|S_i\right|^2\sum_{r=\pm1}\cosh\left(R\right)+\sum_{r=\pm1}\sinh\left(R\right)\Re\left(\chi_\theta\left(\theta,\ n,\ i,\ r\right)\right)\right)\right]^2\cdot$$
$$\left[D\left(\theta,\ G_C,\ X_n,\ \Omega_i,\sigma_W^2\right)\right]^{-2}\Bigg\} \tag{5.9}$$

where we have set, for the sake of readibility:

$$\alpha=\frac{|G_C S_i|}{\sigma_W}$$

$$R\stackrel{\text{def}}{=}\Re\left(\frac{1}{\sigma_W^2}\chi\left(\theta,\ G_C,\ n,\ i,\ r\right)\right)$$

$$D\left(\theta,\ G_C,\ X_n,\ \Omega_i,\sigma_W^2\right)=\sum_{S_i\in\Omega_1}e^{-\frac{|G_C S_i|^2}{2\sigma_W^2}}\sum_{r=\pm1}\cosh\left(\Re\left(\frac{1}{\sigma_W^2}\chi\left(\theta,\ G_C,\ n,\ i,\ r\right)\right)\right)$$

The three expressions in (5.7)-(5.9) can be compactly rewritten as follows:

$$\frac{\partial^2}{\partial\gamma\,\partial\delta}\ln p\left(\mathbf{X}|\theta,\ G_C\right)=\sum_{n=0}^{N-1}\Lambda_{\gamma,\delta}\left(X_n\right)\qquad\text{for }\gamma,\delta\in\{\theta,G_C\}$$

where the $\Lambda_{\gamma,\delta}\left(X_n\right)$ represents the expressions inside of the curly braces. Stemming from this positions we can write the expectation of, for instance, (5.7) as:

$$\mathrm{E}\left\{\frac{\partial^2}{\partial\theta^2}\ln p\left(\mathbf{X}|\theta,\ G_C\right)\right\}=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\sum_{n=0}^{N-1}\Lambda_{\theta,\theta}\left(X_n\right)\prod_{k=0}^{N-1}p\left(X_k|\theta,\ G_C\right)d\left|X_k\right| \tag{5.10}$$

where we denoted $d|X_n| = d\Re(X_n)\, d\Im(X_n)$. Each term $\Lambda_{\theta,\theta}(X_j)$ in the summation in (5.10) depends only on the actual sample X_j and not on the remaining $X_n,\ \ n \neq j$, so that we can write:

$$\mathrm{E}\left\{\frac{\partial^2}{\partial\theta^2}\ln p(\mathbf{X}|\theta,\ G_C)\right\} =$$
$$\int_{-\infty}^{\infty}\Lambda_{\theta,\theta}(X_j)\, p(X_j|\theta,\ G_C)\, d|X_j| \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\prod_{\substack{k=0\\k\neq j}}^{N-1} p(X_k|\theta,\ G_C)\, d|X_k| +$$
$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\sum_{\substack{n=0\\n\neq j}}^{N-1}\Lambda_{\theta,\theta}(X_n)\prod_{k=0}^{N-1} p(X_k|\theta,\ G_C)\, d|X_k|$$

Since we have that

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\prod_{\substack{k=0\\k\neq j}}^{N-1} p(X_k|\theta,\ G_C)\, d|X_k| = 1$$

we can rewrite the expectation in (5.10) as:

$$\mathrm{E}\left\{\frac{\partial^2}{\partial\theta^2}\ln p(\mathbf{X}|\theta,\ G_C)\right\} =$$
$$\sum_{n=0}^{N-1}\int_{-\infty}^{\infty}\Lambda_{\theta,\theta}(X_n)\, p(X_n|\theta,\ G_C)\, d|X_n| =$$
$$N\cdot\mathrm{E}\{\Lambda_{\theta,\theta}(X_n)\}$$

The same result holds also for the other second partial derivatives in (5.8) and (5.9) so that we can write:

$$\mathbf{F}(\theta,\ G_C) = -N\cdot\begin{bmatrix}\mathrm{E}\{\Lambda_{\theta,\theta}(X_n)\} & \mathrm{E}\{\Lambda_{\theta,G_C}(X_n)\}\\ \\ \mathrm{E}\{\Lambda_{\theta,G_C}(X_n)\} & \mathrm{E}\{\Lambda_{G_C,G_C}(X_n)\}\end{bmatrix} \tag{5.11}$$

By defining

$$\mathrm{L}_{\gamma,\delta}\left(\sigma_W^2\right) = \mathrm{E}\{\Lambda_{\gamma,\delta}(X_n)\} \qquad \text{for } \gamma,\delta\in\{\theta, G_C\}$$

we can write:

$$\mathrm{CRB}\left(\hat{\theta}\right) = \frac{\mathrm{L}_{G_C,G_C}(\sigma_W^2)}{N\cdot\left(\mathrm{L}_{G_C,G_C}(\sigma_W^2)\,\mathrm{L}_{\theta,\theta}(\sigma_W^2) - \mathrm{L}_{\theta,G_C}(\sigma_W^2)^2\right)} \tag{5.12}$$

In order to evaluate the $\mathrm{L}_{\gamma,\delta}$'s we resort to a MonteCarlo integration procedure (49). More specifically we produce a large sample of Gaussian noise at different values of SNR, and average the values of the $\Lambda_{\gamma,\delta}(X_n)$'s over these samples. For SNR $\leq$ 20dB the $\Lambda_{\gamma,\delta}(X_n)$'s present spikes too narrow to be treated numerically. Yet, for SNR $\geq$ 20dB the CRB in (5.12) is well approximated by the modified CRB as in (48). The results pertaining this analysis are reported in the performance plots for the phase offset estimator described in Chapter 3 (Figs 3.14-3.19).

5.3 Cramér Rao Lower Bound for joint Frequency and Gain Estimation

In this Section we derive the CRB for joint frequency and gain estimation. We apply much of the derivation for the phase CRB described in the previous Section. The observation model is the same as in 5.1 except for the presence of frequency offset.

$$X_n = G_C\, e^{j(\theta+2\pi f_e n)}\, S_n + W_n \tag{5.13}$$

Following the same steps as in the previous Section we can write the conditional pdf of the received samples as:

$$p(\mathbf{X}|\theta,\ G_C,\ f_e) = \prod_{n=0}^{N-1}\left\{\frac{1}{2\pi M\sigma_W^2}e^{-\frac{|X_n|^2}{2\sigma_W^2}}\sum_{S_i\in\Omega_1}e^{-\frac{|G_C S_i|^2}{2\sigma_W^2}}\cdot\sum_{r=\pm 1}\cosh\left[\Re\left(\frac{1}{\sigma_W^2}\kappa(\theta,\ G_C,\ f_e,\ n,\ i,\ r)\right)\right]\right\} \tag{5.14}$$

where we compactly denoted:

$$\kappa(\theta,\ G_C,\ f_e,\ n,\ i,\ r) \stackrel{\text{def}}{=} |G_C S_i|\, X_n e^{-j\left(\theta+2\pi f_e n+r\arctan\left(\frac{\Im(S_i)}{\Re(S_i)}\right)\right)} \tag{5.15}$$

By comapring the expression in (5.15) with the expression in (5.6) we come up with the following relations:

$$\frac{\partial}{\partial\theta}\Re\left[\kappa\left(\theta,\ G_C,\ f_e,\ n,\ i,\ r\right)\right]=\Im\left[\kappa\left(\theta,\ G_C,\ f_e,\ n,\ i,\ r\right)\right]$$

$$\frac{\partial}{\partial\theta}\Im\left[\kappa\left(\theta,\ G_C,\ f_e,\ n,\ i,\ r\right)\right]=-\Re\left[\kappa\left(\theta,\ G_C,\ f_e,\ n,\ i,\ r\right)\right]$$

$$\frac{\partial}{\partial f_e}\Re\left[\kappa\left(\theta,\ G_C,\ f_e,\ n,\ i,\ r\right)\right]=2\pi n\ \Im\left[\kappa\left(\theta,\ G_C,\ f_e,\ n,\ i,\ r\right)\right]$$

$$\frac{\partial}{\partial f_e}\Im\left[\kappa\left(\theta,\ G_C,\ f_e,\ n,\ i,\ r\right)\right]=-2\pi n\ \Re\left[\kappa\left(\theta,\ G_C,\ f_e,\ n,\ i,\ r\right)\right]$$

$$\frac{\partial}{\partial G_C}\Re\left[\kappa\left(\theta,\ G_C,\ f_e,\ n,\ i,\ r\right)\right]=\Re\left[\frac{\kappa\left(\theta,\ G_C,\ f_e,\ n,\ i,\ r\right)}{G_C}\right]=\Re\left[\kappa_{\theta,f_e}\left(\theta,\ f_e,\ n,\ i,\ r\right)\right]$$

$$\frac{\partial}{\partial G_C}\Im\left[\kappa\left(\theta,\ G_C,\ f_e,\ n,\ i,\ r\right)\right]=\Im\left[\frac{\kappa\left(\theta,\ G_C,\ f_e,\ n,\ i,\ r\right)}{G_C}\right]=\Im\left[\kappa_{\theta,f_e}\left(\theta,\ f_e,\ n,\ i,\ r\right)\right]$$

Stemming from these relations we can write:

$$\frac{\partial^2}{\partial\theta^2}\ln p\left(\mathbf{X}|\theta,\ G_C,\ f_e\right)=\sum_{n=0}^{N-1}\Lambda_{\theta,\theta}\left(X_n\right) \tag{5.16}$$

$$\frac{\partial^2}{\partial G_C^2}\ln p\left(\mathbf{X}|\theta,\ G_C,\ f_e\right)=\sum_{n=0}^{N-1}\Lambda_{G_C,G_C}\left(X_n\right) \tag{5.17}$$

$$\frac{\partial^2}{\partial f_e^2}\ln p\left(\mathbf{X}|\theta,\ G_C,\ f_e\right)=4\pi^2\sum_{n=0}^{N-1}n^2\ \Lambda_{\theta,\theta}\left(X_n\right) \tag{5.18}$$

$$\frac{\partial^2}{\partial\theta\partial f_e}\ln p\left(\mathbf{X}|\theta,\ G_C,\ f_e\right)=2\pi\sum_{n=0}^{N-1}n\ \Lambda_{\theta,\theta}\left(X_n\right) \tag{5.19}$$

$$\frac{\partial^2}{\partial G_C\partial f_e}\ln p\left(\mathbf{X}|\theta,\ G_C,\ f_e\right)=2\pi\sum_{n=0}^{N-1}n\ \Lambda_{G_C,\theta}\left(X_n\right) \tag{5.20}$$

$$\frac{\partial^2}{\partial G_C\partial\theta}\ln p\left(\mathbf{X}|\theta,\ G_C,\ f_e\right)=\sum_{n=0}^{N-1}\Lambda_{G_C,\theta}\left(X_n\right) \tag{5.21}$$

with the care of cosidering $\kappa(\theta,\ G_C,\ f_e,\ n,\ i,\ r)$, instead of $\chi(\theta,\ G_C,\ n,\ i,\ r)$ in the definition of the terms $\Lambda_{\gamma,\delta}$ introduced in the previous Section.

Again the expectations can be taken follwing the guidelines depicted for the case of phase offset estimation, and resorting to Monte Carlo integration technique to numerically evaluate the expectations involved in evaluating the terms in the Fisher Information Matrix.

The results of this analysis are found in Figs. 4.4 - 4.9 reporting the accuracy of the frequency offset estimator introduced in Chapter. 4.

Chapter 6

Conclusion

In this thesis a Generalized Method of Moments (GMM) approach for (cyclic) shift parameter estimation has been introduced. We showed how the shift nature of the *estimandum* is suitably exploited to perform the minimization problem comprised in the general GMM framework via a DFT-based approach. This approach, by itself suitable only when a location parameter is to be estimated, can be correctly employed also when a transformation can be found, such to map the parameter estimation problem into a shift estimation one. An example of this scenario is, for instance, the gain estimation in the Mellin transformed domain. Besides, we have discussed the condition under which the GMM provides a ML estimation, proving that, when the observations are multinomially distributed, a MLE is obtained by the same estimation rule of the so-called unweighted GMM with a significantly reduced computational complexity.

As an application we introduced a novel gain control free, not data aided, phase and frequency offset estimation technique for QAM signaling. The estimators stem from the observation of the peculiar features that the constellation diagrams of the received samples exhibits in presence of phase and/or frequency offset. An image processing like analysis of such features helps defining a statistic so to drive the estimation. More specifically the statistic is based on the sample measurement of the tomographic projection of the bidimensional pdf of the received data, which we have called *Constellation Phase Signature* (CPS).

The phase-offset causes a cyclic shift of the CPS, and, thus, we have rephrased the phase-offset estimation problem rather as a cyclic shift estimation problem,

and we applied the GMM framework for shift parameter estimation herein presented. Moreover the samples of the measured CPS, under a particular setting, can be proved to be multinomially distributed, thus enabling a reduced complexity ML phase offset estimation.

The frequency-offset is instead proved to cause a filtering over the estimated CPS, due to the time varying phase drift each received sample is subject to. A suitable frequency compensation removes the time varying phase drift and restores the original structure of the CPS. Hence, the frequency estimation problem has been reformulated in terms of maximization of the cross-correlation between the sample CPS and the analytically known CPS for zero frequency offset.

We have carried out a theoretical analysis concerning with asymptotical performance, which has been assessed by numerical simulations. We compared the CPS based estimators with selected existing blind estimators and with the Cramér Rao Lower Bound (CRB), which we explicitely evaluated for the reference case of unknown gain factor. We have shown that the variance of the CPS based estimator versus SNR has the same slope of CRB, and, in the case of optimally weighted GMM and for the reduced complexity ML estimators, it approaches the CRB for a wide range of SNR. Moreover, the CPS based estimators outperform selected existing blind estimators for medium to high values of SNR and for dense constellations.

Chapter 7

Appendices

A Variances-Covariances of the Weighted Histogram

Let us here rewrite the sample estimates defined in (2.3), in the case of K equispaced points $\xi_k = (2k+1)\pi/K$, for $k = 0, \ldots, K-1$:

$$\hat{f}_k = \frac{K}{2\pi\, N} \sum_{n=1}^{N} h\left(z_n^{(1)}, \ldots, z_n^{(l-1)}\right) \operatorname{rect}\left(\frac{z_n^{(l)} - \xi_k}{2\pi/K}\right) \tag{A.1}$$

where the observations $z_n^{(1)}, \ldots, z_n^{(l)}$ are drawn in a statistically independent fashion for $n = 1, N$.

For the sake of compactness, let us indicate the $l-1$ random variables $z^{(1)} \cdots z^{(l-1)}$ as a single, $(l-1)$-dimensional, random variable ζ. Then, as far as the first order moment of $\hat{f}_k$ is concerned, we have:

$$\begin{aligned} \mathrm{E}\left\{\hat{f}_k\right\} = \frac{K}{2\pi} \int_{\xi_k - \pi/K}^{\xi_k + \pi/K} p_{\mathrm{z}^{(l)}|\alpha}\left(z^{(l)}|\alpha\right) dz^{(l)} \\ \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} h\left(\zeta\right) p_{\zeta|\mathrm{z}^{(l)},\alpha}\left(\zeta|z^{(l)},\alpha\right)\, d\zeta \end{aligned} \tag{A.2}$$

where $d\zeta$ stands for $dz^{(1)} \cdots dz^{(l-1)}$. The $(l-1)$-dimensional inner integral in (A.2) is the expectaton of the function $h\left(\zeta\right)$ conditioned to $z^{(l)}$. Then, let us

shortly denote conditional moments of $h(\zeta)$ as:

$$\mu^{(k)}_{h|z^{(l)}} \stackrel{\text{def}}{=} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} h^k(\zeta)\, p_{\zeta|\mathrm{z}^{(l)},\alpha}\left(\zeta|z^{(l)},\alpha\right)\, d\zeta$$

in order to rewrite (A.2) as follows:

$$\begin{aligned} \mathrm{E}\left\{\hat{f}_k\right\} &= \frac{K}{2\pi}\int_{\xi_k-\pi/K}^{\xi_k+\pi/K} \mu^{(1)}_{h|z^{(l)}}\, p_{\mathrm{z}^{(l)}|\alpha}\left(z^{(l)}|\alpha\right)\, dz^{(l)} \\ &= \frac{K}{2\pi}\,\mathrm{E}\left\{\mu^{(1)}_{h|z^{(l)}}\,\middle|\,\mathcal{E}_k\right\} \end{aligned} \tag{A.3}$$

where $\mathcal{E}_k = \left\{|z^{(l)} - \xi_k| < 2\pi/K\right\}$ form a complete class of mutually exclusive events.

In the case $h(\cdot) = 1$, we are faced with the classical (unweighted) marginal histogram wrt the random variable $z^{(l)}$, and (A.3) rediscovers the well know property of unbiasness of histograms: $2\pi\mathrm{E}\left\{\hat{f}_k\right\}/K = \mathrm{Prob}\{\mathcal{E}_k\}$. For a general nonlinearity $h(\cdot)$, recalling the definition (2.2), we have:

$$\frac{K}{2\pi}\,\mathrm{E}\left\{\mu^{(i)}_{h|z^{(l)}}\,\middle|\,\mathcal{E}_k\right\} = f\left(\xi_k - \alpha\right)$$

Hence, (A.3) expresses the unbiasness of weighted histograms:

$$\mathrm{E}\left\{\hat{f}_k\right\} = f\left(\xi_k - \alpha\right) \tag{A.4}$$

As far as the variances-covariances are concerned, we have:

$$
\begin{aligned}
\mathrm{Cov}\left\{\hat{f}_k, \hat{f}_q\right\} = \frac{K^2}{4\pi^2\, N^2} \sum_{n=1}^{N} & \\
\mathrm{Cov}\left\{ h\left(\zeta_n\right) \mathrm{rect}\left(\frac{z_n^{(l)} - \xi_k}{2\pi/K}\right), h\left(\zeta_n\right) \mathrm{rect}\left(\frac{z_n^{(l)} - \xi_q}{2\pi/K}\right)\right\} & \\
+ \frac{K^2}{4\pi^2\, N^2} \sum_{n=1}^{N} \sum_{\substack{m=1 \\ m\neq n}}^{N} & \\
\underbrace{\mathrm{Cov}\left\{ h\left(\zeta_n\right) \mathrm{rect}\left(\frac{z_n^{(l)} - \xi_k}{2\pi/K}\right), h\left(\zeta_m\right) \mathrm{rect}\left(\frac{z_m^{(l)} - \xi_q}{2\pi/K}\right)\right\}}_{0 \text{ (statistically independent random variables)}} & \\
= \frac{K^2}{4\pi^2\, N}\Bigg[\mathrm{Var}\left\{ h\left(\zeta\right) \mathrm{rect}\left(\frac{z^{(l)} - \xi_k}{2\pi/K}\right)\right\} \cdot \delta_{k,q} & \\
- \mathrm{E}\left\{ h\left(\zeta\right) \mathrm{rect}\left(\frac{z^{(l)} - \xi_k}{2\pi/K}\right)\right\} & \\
\cdot \mathrm{E}\left\{ h\left(\zeta\right) \mathrm{rect}\left(\frac{z^{(l)} - \xi_q}{2\pi/K}\right)\right\} \cdot \left(1 - \delta_{k,q}\right)\Bigg] & \\
= \frac{K^2}{4\pi^2\, N}\Big[\mathrm{E}\left\{ \mu^{(2)}_{h|z^{(l)}} \,\middle|\, \mathcal{E}_k \right\} \delta_{k,q} & \\
- \mathrm{E}\left\{ \mu^{(1)}_{h|z^{(l)}} \,\middle|\, \mathcal{E}_k \right\} \cdot \mathrm{E}\left\{ \mu^{(1)}_{h|z^{(l)}} \,\middle|\, \mathcal{E}_q \right\} \Big] &
\end{aligned}
\tag{A.5}
$$

being $\delta_{k,q}$ the Kronecker delta, *i.e.* it is equal to 1 for $k = q$ and 0 elsewhere.

The evaluation of the conditional moments $\mathrm{E}\left\{ \mu^{(k)}_{h|z^{(l)}} \,\middle|\, \mathcal{E}_k \right\}$, for $k = 1, 2$, is carried out on the basis of the probabilistic description of the specific problem at hand. For instance, in the reference case of phse acquisition, in App.E and F analytical expressions of the first and second order conditional moments are respectively reported.

B GMM Estimator Statistical Analysis

For simplicity, here and in the following we will drop out the dependence on α in the optimal weight matrix by using the notation $\mathbf{W}_0 \stackrel{\text{def}}{=} \mathbf{W}^{(o)}$.

As far as the first order moment is concerned, since $\mathrm{E}\{\hat{\mathbf{f}}\} = \mathbf{f}(\alpha)$ we have:

$$\mathrm{E}\left\{\mathcal{J}\left(\alpha; \mathbf{W}_0\right)\right\} = \left[\mathbf{f}(\alpha) - \frac{1}{2} \cdot \mathbf{f}(\alpha)\right]^{\mathrm{T}} \cdot \mathbf{W}_0 \cdot \mathbf{f}(\alpha)$$

The variances-covariances are evaluated as follows:

$$\begin{aligned} & N \cdot \mathrm{Cov}\left\{\mathcal{J}\left(\alpha_1; \mathbf{W}_0\right), \mathcal{J}\left(\alpha_2; \mathbf{W}_0\right)\right\} \\ & \quad = N \cdot \mathbf{f}(\alpha_1)^{\mathrm{T}} \cdot \mathbf{W}_0 \cdot \mathrm{Cov}\left\{\hat{\mathbf{f}}, \hat{\mathbf{f}}^{\mathrm{T}}\right\} \cdot \mathbf{W}_0 \cdot \mathbf{f}(\alpha_2) \\ & \quad = N \cdot \mathbf{f}(\alpha_1)^{\mathrm{T}} \cdot \mathbf{W}_0 \cdot \mathbf{\Omega}(\theta) \cdot \mathbf{W}_0 \cdot \mathbf{f}(\alpha_2) \\ & \quad = \mathbf{f}(\alpha_1)^{\mathrm{T}} \cdot \mathbf{W}_0 \cdot \mathbf{f}(\alpha_2) \end{aligned} \tag{B.1}$$

For the expression of $\mathrm{Cov}\left\{\hat{\mathbf{f}}, \hat{\mathbf{f}}^{\mathrm{T}}\right\}$, refer to App.A.

C Reduced Complexity ML Estimator Statistical Analysis

Since $\mathrm{E}\{\hat{\mathbf{f}}\} = \mathbf{f}(\alpha)$ we have:

$$\mathrm{E}\left\{\mathcal{C}(\alpha)\right\} = \mathbf{f}(\alpha)^{\mathrm{T}} \cdot \tilde{\mathbf{f}}(\alpha)$$

The variances-covariances are evaluated as follows:

$$\begin{aligned} & N \cdot \mathrm{Cov}\left\{\mathcal{C}(\alpha_1), \mathcal{C}(\alpha_2)\right\} \\ & \quad = N \cdot \tilde{\mathbf{f}}(\alpha_1)^{\mathrm{T}} \cdot \mathrm{Cov}\left\{\hat{\mathbf{f}}, \hat{\mathbf{f}}^{\mathrm{T}}\right\} \cdot \tilde{\mathbf{f}}(\alpha_2) \end{aligned} \tag{C.1}$$

For the expression of $\mathrm{Cov}\left\{\hat{\mathbf{f}}, \hat{\mathbf{f}}^{\mathrm{T}}\right\}$, refer to App.A.

D On The Optimal Weight Matrix $\mathbf{W}_0$

The computation of the optimal weight matrix $\mathbf{W}_0 = \Omega(\alpha)^{-1}/N$ requires the inversion of the covariance matrix of the measurements $\Omega(\alpha) \stackrel{\text{def}}{=} \mathrm{E}\{(\hat{\mathbf{f}} - \mathbf{f}(\alpha))(\hat{\mathbf{f}} - \mathbf{f}(\alpha))^{\mathrm{T}}\}$, which, according to (A.5), can be expressed as follows:

$$N \cdot \Omega(\alpha) = \mathbf{\Lambda}(\alpha) - \mathbf{f}(\alpha) \cdot \mathbf{f}(\alpha)^{\mathrm{T}}$$

where

$$\mathbf{\Lambda}(\alpha) = \mathrm{diag}\left\{g_1^2, \ldots, g_K^2\right\}$$

having defined $g_k^2 \stackrel{\text{def}}{=} \frac{K}{2\pi}\mathrm{E}\left\{\mu_{h|z^{(l)}}^{(1)} \middle| \mathcal{E}_k\right\}$. Using the Woodbury's identity (22), we obtain:

$$\mathbf{W}_0 = \mathbf{\Lambda}(\alpha)^{-1} + \frac{\mathbf{\Lambda}(\alpha)^{-1} \cdot \mathbf{f}(\alpha) \cdot \mathbf{f}(\alpha)^{\mathrm{T}} \cdot \mathbf{\Lambda}(\alpha)^{-1}}{1 - \mathbf{f}(\alpha)^{\mathrm{T}} \cdot \mathbf{\Lambda}(\alpha)^{-1} \cdot \mathbf{f}(\alpha)} \tag{D.1}$$

Hence, the elements of the matrix $\mathbf{W}_0$ take the following form:

$$||\mathbf{W}_0||_{k,q} = \frac{\delta_{k,q}}{g_k^2} + \left(1 - \sum_{m=1}^{K} \frac{\hat{f}_m^2}{g_m^2}\right)^{-1} \frac{\hat{f}_k}{g_k^2} \, \frac{\hat{f}_q}{g_q^2}$$

It is worth noting that the computations needed to evaluate (B.1) and (C.1) are significantly reduced when the optimal weight matrix $\mathbf{W}_0$ is expressed in the form (D.1).

E Expectation of the Sample CPS and Analytical Evaluation of the Zero Phase-Offset MWTP

We start by evaluating the expectation of the sample CPS. At this aim, let us rewrite the signal model in (3.1) as $X[n] = G_C\,(e^{j\theta}\,S[n] + W'[n])$, being $W'[n] = W[n]/G_C$ a realization of a circularly complex Gaussian stationary noise process with variance $\sigma_W^2/G_C^2 = 1/\mathrm{SNR}$. At an assigned SNR value, the actual value of G_C influences the nonlinear function $Y[n]$ defined in (3.2), as well as the CPS introduced in (3.8), only by a constant factor G_C^P, which in turn does not affect the phase-offset estimate defined in (3.13). Hence, without loss of generality, in the following we will proceed assuming $G_C = 1$; let us remark that, within this assumption, the noise variance σ_W^2 assumes the physical meaning of Noise-to-Signal Ratio.

For $k = 0, \ldots, K-1$, dropping out the time index n for simplicity, we have:

$$\begin{aligned}\lim_{K\to\infty} \mathrm{E}\left\{\hat{f}^{(\mathcal{A},\theta,P)}\,(\psi_k)\right\} &= \lim_{K\to\infty} \mathrm{E}\left\{\frac{K}{2\pi}\cdot|Y|d_K^{(k)}\,(Y)\right\} \\ &= \lim_{K\to\infty}\int \frac{K}{2\pi}\cdot|Y|d_K^{(k)}\,(Y)\;\; p_Y(Y)\,dY \\ &= \lim_{K\to\infty}\int\limits_{\arg Y\in I_k} \frac{K}{2\pi}\cdot|Y|\;p_Y(Y)\,dY\end{aligned} \tag{E.1}$$

We recall that the discrete argument $k \in [0, K-1]$ indexes the angular window $I_K^{(k)}$, defined in (3.6) and here repeated for the sake of readability:

$$I_K^{(k)} \stackrel{\mathrm{def}}{=} \left[k\cdot\frac{2\pi}{K}, (k+1)\cdot\frac{2\pi}{K}\right), \quad \text{for } k = 0, \ldots, K-1 \tag{E.2}$$

In the limit $K \to \infty$, the interval $I_K^{(k)}$ becomes infinitesimal, $2\pi/K \to d\psi$, and the continuous phase argument $2\pi k/K \to \psi$ spans the range $[0, 2\pi)$ in a periodic, modulo 2π, fashion.

In polar coordinates $Y = re^{j\varphi}$, the last integral in (E.1) is written as follows:

$$\begin{aligned}&\lim_{K\to\infty} \mathrm{E}\left\{\hat{f}^{(\mathcal{A},\theta,P)}\,(\psi_k)\right\} \\ &\qquad = \lim_{K\to\infty}\int_0^{+\infty}\int_{2\pi k/K}^{2\pi(k+1)/K} \frac{K}{2\pi}\,r\,p_{R,\Phi}\,(r,\varphi;\theta)\,dr d\varphi\end{aligned}$$

In the limit $K \to \infty$, the integral with respect to φ tends to a line integral defined on the straight line where the phase φ equals ψ, *i.e.*:

$$\begin{aligned}\lim_{K\to\infty} \mathrm{E}\left\{\hat{f}^{(\mathcal{A},\theta,P)}\left(\psi_k\right)\right\} &= \int_0^{+\infty}\int_0^{2\pi} r\,\delta(\varphi-\psi)\,p_{R,\Phi}\left(r,\varphi;\theta\right)dr d\varphi \\ &= \int_0^{+\infty} r\,p_{R,\Phi}\left(r,\psi;\theta\right)dr \\ &= g_{\Phi}^{(\mathcal{A},\theta,P)}\left(\psi_k\right)\end{aligned}$$

and this shows that the sample CPS $\hat{f}^{(\mathcal{A},\theta,P)}\left(\psi_k\right)$ is an unbiased estimator of the MWTP $g_{\Phi}^{(\mathcal{A},\theta,P)}\left(\psi_k\right)$.

In order to proceed toward the evaluation of the zero phase-offset MWTP $g_{\Phi}^{(\mathcal{A},0,P)}\left(\psi_k\right)$, let us recall that $Y = |X|^P e^{j4\cdot\arg X}$. Setting $X = \gamma\, e^{j\phi}$, we can rewrite the previous integral as follows:

$$g_{\Phi}^{(\mathcal{A},0,P)}\left(\psi_k\right) = 4\int_0^{\infty} \gamma^P p_{\Gamma,\Phi}(\gamma,\psi_k/4;0)\,d\gamma$$

Using the polar representation $S_m = \gamma_m\, e^{j\phi_m}$. for the generic constellation symbol, for an equiprobable M-constellation the pdf of the received signal $X = \gamma\, e^{j\phi}$ is:

$$\begin{aligned}p_{\Gamma,\Phi}(\gamma,\phi;0) = \sum_{m=0}^{M-1} \frac{\gamma}{M\,\pi\,\sigma_W^2} \exp\left(-\frac{\left(\gamma\cos\phi-\gamma_m\cos\phi_m\right)^2}{\sigma_W^2}\right) \\ \cdot \exp\left(-\frac{\left(\gamma\sin\phi-\gamma_m\sin\phi_m\right)^2}{\sigma_W^2}\right)\end{aligned}$$

After some calculation we obtain:

$$
\begin{aligned}
g_{\Phi}^{(\mathcal{A},0,P)}(\psi_k) &= \frac{4}{M}\sum_{m=0}^{M-1}\frac{1}{\pi\,\sigma_W^2}\exp\left(-\frac{\gamma_m^2\sin^2(\psi_k/4-\phi_m)}{\sigma_W^2}\right)\\
&\quad\cdot\int_{-\gamma_m\cos(\psi_k/4-\phi_m)}^{\infty}\left[\gamma+\gamma_m\cos(\psi_k/4-\phi_m)\right]^{P+1}\exp\left(-\frac{\gamma 2}{\sigma_W^2}\right)d\gamma\\
&= \frac{4}{M}\sum_{m=0}^{M-1}\frac{1}{\pi\,\sigma_W^2}\exp\left(-\frac{\gamma_m^2\sin^2(\psi_k/4-\phi_m)}{\sigma_W^2}\right)\\
&\quad\cdot\sum_{p=0}^{P+1}\binom{P+1}{p}\left(\gamma_m\cos(\psi_k/4-\phi_m)\right)^{P+1-p}\\
&\quad\cdot\int_{-\gamma_m\cos(\psi_k/4-\phi_m)}^{\infty}\gamma^p\exp\left(-\frac{\gamma^2}{\sigma_W^2}\right)d\gamma\\
&= \frac{4}{M}\sum_{m=0}^{M-1}\frac{1}{\pi\,\sigma_W^2}\exp\left(-\frac{\gamma_m^2\sin^2(\psi_k/4-\phi_m)}{\sigma_W^2}\right)\\
&\quad\cdot\sum_{p=0}^{P+1}\binom{P+1}{p}\left(\gamma_m\cos(\psi_k/4-\phi_m)\right)^{P+1-p}\\
&\quad\cdot U^{(p)}(\psi_k,\gamma_m,\phi_m,\sigma_W)
\end{aligned}
$$

where we have set

$$
U^{(p)}(\psi_k,\gamma_m,\phi_m,\sigma_W)\overset{\text{def}}{=}\int_{-\gamma_m\cos(\psi_k/4-\phi_m)}^{\infty}\gamma^p\exp\left(-\frac{\gamma^2}{\sigma_W^2}\right)d\gamma \tag{E.3}
$$

We have analytically evaluated the integrals in (E.3), and the functions $U^{(k)}(\psi_k, \gamma_m, \phi_m, \sigma_W)$ take the following form:

$$
\begin{aligned}
U^{(0)} &= \frac{\sqrt{\pi}}{2}\sigma_W \,\mathrm{Erfc}\left(\frac{\gamma_m \cos(\psi_k/4-\phi_m)}{\sigma_W}\right) \\
U^{(1)} &= \frac{1}{2}\sigma_W^2 \;\exp\left(-\frac{(\gamma_m \cos(\psi_k/4-\phi_m))^2}{\sigma_W^2}\right) \\
U^{(2)} &= \frac{1}{4}\sigma_W^2\Big[\exp\left(-\frac{(\gamma_m \cos(\psi_k/4-\phi_m))^2}{\sigma_W^2}\right) 2\gamma_m \cos(\psi_k/4-\phi_m) \\
&\quad +\sqrt{\pi}\sigma_W - \sqrt{\pi}\sigma_W \,\mathrm{Erf}\left(\frac{\gamma_m \cos(\psi_k/4-\phi_m)}{\sigma_W}\right)\Big] \\
U^{(3)} &= \frac{1}{2}\sigma_W^2 \;\exp\left(-\frac{(\gamma_m \cos(\psi_k/4-\phi_m))^2}{\sigma_W^2}\right) \\
&\quad \cdot \left((\gamma_m \cos(\psi_k/4-\phi_m))^2 + \sigma_W^2\right) \\
U^{(4)} &= \frac{1}{8}\sigma_W^2\left[\exp\left(-\frac{(\gamma_m \cos(\psi_k/4-\phi_m))^2}{\sigma_W^2}\right)\right. \\
&\quad \cdot (4(\gamma_m \cos(\psi_k/4-\phi_m))^3 + 6\gamma_m \cos(\psi_k/4-\phi_m)\sigma_W^2) \\
&\quad \left. + 3\sqrt{\pi}\sigma_W^3 - 3\sqrt{\pi}\sigma_W^3 \,\mathrm{Erf}\left(\frac{\gamma_m \cos(\psi_k/4-\phi_m)}{\sigma_W}\right)\right] \\
U^{(5)} &= \frac{1}{2}\sigma_W^2 \;\exp\left(-\frac{(\gamma_m \cos(\psi_k/4-\phi_m))^2}{\sigma_W^2}\right) \\
&\quad \cdot \left[(\gamma_m \cos(\psi_k/4-\phi_m))^4 2\sigma_W^2 (\gamma_m \cos(\psi_k/4-\phi_m))^2 + 2\sigma_W^4\right]
\end{aligned}
$$

The expression of the theoretical CPS can be clearly recognized as an example of the first order conditional moment of weighted histogram as generally introduced and defined in App.A.

F Analytical Evaluation of the Variances/Covariances of the Sample CPS

As far as the expected value of the sample CPS $\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)$ is concerned, we have:

$$\mathrm{E}\left\{\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)\right\} = \mathrm{E}\left\{|Y|\cdot d_K^{(k)}(Y)\right\}$$

For the variances we have:

$$\begin{aligned}
&\mathrm{Var}\left\{\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)\right\} \stackrel{\mathrm{def}}{=} \mathrm{E}\left\{\left(\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)\right)^2\right\} - \left(\mathrm{E}\left\{\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)\right\}\right)^2 \\
&= \mathrm{E}\left\{\frac{1}{N^2}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1}|Y[n]|\cdot d_K^{(k)}(Y[n])\,|Y[m]|\cdot d_K^{(k)}(Y[m])\right\} - \left(\mathrm{E}\left\{|Y|\cdot d_K^{(k)}(Y)\right\}\right)^2 \\
&= \frac{1}{N^2}\left[N\cdot\mathrm{E}\left\{|Y|^2\cdot d_K^{(k)}(Y)\right\} + (N^2-N)\left(\mathrm{E}\left\{|Y|\cdot d_K^{(k)}(Y)\right\}\right)^2\right] - \left(\mathrm{E}\left\{|Y|\cdot d_K^{(k)}(Y)\right\}\right)^2 \\
&= \frac{1}{N}\left[\mathrm{E}\left\{|Y|^2\cdot d_K^{(k)}(Y)\right\} - \left(\mathrm{E}\left\{|Y|\cdot d_K^{(k)}(Y)\right\}\right)^2\right]
\end{aligned}$$

For the covariances we have:

$$\begin{aligned}
&\mathrm{Cov}\left\{\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k),\hat{f}^{(\mathcal{A},\theta,P)}(\psi_l)\right\} \stackrel{\mathrm{def}}{=} \\
&\mathrm{E}\left\{\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)\,\hat{f}^{(\mathcal{A},\theta,P)}(\psi_l)\right\} - \mathrm{E}\left\{\hat{f}^{(\mathcal{A},\theta,P)}(\psi_k)\right\}\mathrm{E}\left\{\hat{f}^{(\mathcal{A},\theta,P)}(\psi_l)\right\} \\
&= \mathrm{E}\left\{\frac{1}{N^2}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1}|Y[n]|\cdot d_K^{(k)}(Y[n])\,|Y[m]|\cdot d_K^{(l)}(Y[m])\right\} \\
&\quad - \mathrm{E}\left\{|Y|\cdot d_K^{(k)}(Y)\right\}\mathrm{E}\left\{|Y|\cdot d_K^{(l)}(Y)\right\} \\
&= \frac{1}{N^2}\Big(N\cdot\mathrm{E}\left\{|Y|^2\cdot d_K^{(k)}(Y)\cdot d_K^{(l)}(Y)\right\} + (N^2-N)\cdot\mathrm{E}\left\{|Y|\cdot d_K^{(k)}(Y)\right\}\mathrm{E}\left\{|Y|\cdot d_K^{(l)}(Y)\right\}\Big) \\
&\quad - \mathrm{E}\left\{|Y|\cdot d_K^{(k)}(Y)\right\}\mathrm{E}\left\{|Y|\cdot d_K^{(l)}(Y)\right\} \\
&= \frac{1}{N}\left(\mathrm{E}\left\{|Y|^2\cdot d_K^{(k)}(Y)\right\}\delta_{kl} - \mathrm{E}\left\{|Y|\cdot d_K^{(k)}(Y)\right\}\mathrm{E}\left\{|Y|\cdot d_K^{(l)}(Y)\right\}\right)
\end{aligned}$$

where δ_{kl} is the Kronecker delta, *i.e.* $\delta_{kl} = 1$ for $k = l$ and $\delta_{kl} = 0$ for $k \neq l$.

In the limit $K \to \infty$ we obtain

$$\mathrm{E}\left\{\hat{f}^{(\mathcal{A},\theta,P)}\left(\psi_k\right)\right\} = g_{\Phi}^{(\mathcal{A},\theta,P)}\left(\psi_k\right)$$

$$N \cdot \mathrm{Var}\left\{\hat{f}^{(\mathcal{A},\theta,P)}\left(\psi_k\right)\right\} = g_{\Phi}^{(\mathcal{A},\theta,2P)}\left(\psi_k\right) - \left(g_{\Phi}^{(\mathcal{A},\theta,P)}\left(\psi_k\right)\right)^2$$

$$N \cdot \mathrm{Cov}\left\{\hat{f}^{(\mathcal{A},\theta,P)}\left(\psi_k\right), \hat{f}^{(\mathcal{A},\theta,P)}\left(\psi_l\right)\right\} = g_{\Phi}^{(\mathcal{A},\theta,2P)}\left(\psi_k\right)\delta_{kl} - g_{\Phi}^{(\mathcal{A},\theta,P)}\left(\psi_k\right) \cdot g_{\Phi}^{(\mathcal{A},\theta,P)}\left(\psi_l\right)$$

These results particularize the expression of the conditional moments in App.A

G A rough SNR estimator

The bandwidth of CPS is related to the SNR. In fact, since the width of the pulses of the CPS increases as the SNR decreases, the bandwidth is expected to decrease too. We have measured the bandwidth B of the expected CPS for a discrete set of SNR values. Specifically, fixed the constellation $\mathcal{A}$, for each SNR value we have calculated:

$$B(\text{SNR}) = \sqrt{\frac{\sum_{k=-k_{\mathcal{A}}}^{k_{\mathcal{A}}} k^2 \cdot M[k]}{\sum_{k=-k_{\mathcal{A}}}^{k_{\mathcal{A}}} M[k]}} \tag{G.1}$$

being $M[k]$ is the magnitude of the DFT of the expected CPS at that SNR. The index $k_{\mathcal{A}}$, limiting the calculation to low frequencies, controls the possible noise amplification occurring in the summation of (G.1), and it can be chosen accordingly to the particular constellation. We have set $k_{\mathcal{A}} = 8$ for all the here considered *constellations* . For each constellation, the values $B(\text{SNR})$ have been collected in a table; in the estimation, the table is then looked up to obtain the SNR value once $B(\text{SNR})$ is instead calculated using (G.1) with $M[k]$ being the magnitude of the DFT of the sample CPS. It is worth noting that no significant computational load is added to the overall estimation process since the DFT of the sample CPS has been previously calculated to implement in the frequency domain the cyclic cross-correlation in (3.13).

We remark that for constellations *richer* than 256-QAM the bandwidth measurement in (G.1) is very unstable for low SNR values; thus, to obtain the numerical results of Fig.3.13 pertaining to 512-QAM, we have replaced (G.1) with the modified $B(\text{SNR}) = M[0]/\sqrt{\sum_{k=-k_{\mathcal{A}}}^{k_{\mathcal{A}}} M^2[k]}$ when it results $\sum_{k=-k_{\mathcal{A}}}^{k_{\mathcal{A}}} M[k] \ll 2M[0]$, which happens for SNR < 30dB.

H Statistical Analysis of the Frequency-Offset Objective Function

In this Appendix we report the evaluation of the first and second order moments of the objective function $\mathcal{G}(f_c)$, as given in (4.16), in terms of the MWTP of the QAM signals.

In order to evaluate the variance and covariance of $\mathcal{G}(f_c)$, let us rewrite (4.16) as follows:

$$\begin{aligned}\mathcal{G}(f_c) &= \sum_{m=0}^{K-1} \left| A^{(\mathcal{A},(f_c-f_0),\theta)}(m) \right| \cdot \left| G_{\Phi}^{(\mathcal{A},0)}(m) \right| \\ &\simeq \sum_{m=0}^{K-1} a^{(\mathcal{A},(f_c-f_0),0)}(\psi_k) \cdot g_{\Phi}^{(\mathcal{A},0,P)}(\psi_k)\end{aligned} \tag{H.1}$$

where, without loss of generality, we have set $\theta = 0$.

Introducing the zero mean random process

$$\epsilon_k(f_c - f_0) \stackrel{\text{def}}{=} a^{(\mathcal{A},(f_c-f_0),0)}(\psi_k) - \mathrm{E}\left\{ a^{(\mathcal{A},(f_c-f_0),0)}(\psi_k) \right\}$$

where the expectation $\mathrm{E}\left\{ a^{(\mathcal{A},(f_c-f_0),0)}(\psi_k) \right\}$ is given in (4.11), we can rewrite the objective function in (H.1) as follows:

$$\mathcal{G}(f_c) = \sum_{k=0}^{K-1} \left(\mathrm{E}\left\{ a^{(\mathcal{A},(f_c-f_0),0)}(\psi_k) \right\} + \epsilon_k(f_c - f_0) \right) \cdot g_{\Phi}^{(\mathcal{A},0,P)}(\psi_k)$$

The first and second order moments of $\mathcal{G}(\cdot)$ are then evaluated as follows:

$$\mathrm{E}\{\mathcal{G}(f_c)\} = \frac{1}{K} \sum_{k=0}^{K-1} \mathrm{E}\left\{ a^{(\mathcal{A},(f_c-f_0),0)}(\psi_k) \right\} \cdot g_{\Phi}^{(\mathcal{A},0,P)}(\psi_k)$$

$$\begin{aligned}\mathrm{E}\left\{\mathcal{G}^2(f_c)\right\} = \frac{1}{K^2} \sum_{k_1}^{K-1} \sum_{k_2}^{K-1} \Bigg\{ & \mathrm{E}\left\{ a^{(\mathcal{A},(f_c-f_0),0)}(\psi_{k_1}) \right\} \\ & \cdot g_{\Phi}^{(\mathcal{A},0,P)}(\psi_{k_1}) \, \mathrm{E}\left\{ a^{(\mathcal{A},(f_c-f_0),0)}(\psi_{k_2}) \right\} g_{\Phi}^{(\mathcal{A},0,P)}(\psi_{k_2}) \\ & + g_{\Phi}^{(\mathcal{A},0,P)}(\psi_{k_1}) \, g_{\Phi}^{(\mathcal{A},0,P)}(\psi_{k_2}) \, m_{f_c}^{(2)}[k_1, k_2] \Bigg\}\end{aligned}$$

$$\mathrm{E}\left\{\mathcal{G}(f_{c_1})\mathcal{G}(f_{c_2})\right\} = \frac{1}{K^2}\sum_{k_1}^{K-1}\sum_{k_2}^{K-1}\Bigg\{\mathrm{E}\left\{a^{(\mathcal{A},(f_{c_1}-f_0),0)}\left(\psi_{k_1}\right)\right\} \cdot g_{\Phi}^{(\mathcal{A},0,P)}\left(\psi_{k_1}\right)\mathrm{E}\left\{a^{(\mathcal{A},(f_{c_2}-f_0),0)}\left(\psi_{k_2}\right)\right\} g_{\Phi}^{(\mathcal{A},0,P)}\left(\psi_{k_2}\right) + g_{\Phi}^{(\mathcal{A},0,P)}\left(\psi_{k_1}\right) g_{\Phi}^{(\mathcal{A},0,P)}\left(\psi_{k_2}\right) m_{f_{c_1},f_{c_2}}^{(1,1)}[k_1,k_2]\Bigg\}$$

where we have compactly denoted

$$m_{f_c}^{(2)}[k] \stackrel{\text{def}}{=} \mathrm{E}\left\{\left(\epsilon_k(f_c-f_0)\right)^2\right\}$$
$$m_{f_c}^{(1,1)}[k_1,k_2] \stackrel{\text{def}}{=} \mathrm{E}\left\{\epsilon_{k_1}(f_c-f_0)\cdot\epsilon_{k_2}(f_c-f_0)\right\}$$
$$m_{f_{c_1},f_{c_2}}^{(1,1)}[k_1,k_2] \stackrel{\text{def}}{=} \mathrm{E}\left\{\epsilon_{k_1}(f_{c_1}-f_0)\cdot\epsilon_{k_2}(f_{c_2}-f_0)\right\}$$

H–1 Moments of the Error Process

In order to complete the statistical analysis, we carry out the detailed analytical evaluation of the moments of the error process $\epsilon_k(f_c - f_0)$.

The second order moment of the accumulation function $\mathrm{E}\left\{a^{(\mathcal{A},(f_c-f_0),0)}\left(\psi_k\right)^2\right\}$ is evaluated as follows:

$$\begin{aligned}
\mathrm{E}\left\{a^{(\mathcal{A},(f_c-f_0),0)}\left(\psi_k\right)^2\right\} &= \mathrm{E}\Bigg\{\frac{1}{N^2}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1}\left|Y_n^{(f_c)}\right| d_K^{(k)}\left(Y_n^{(f_c)}\right)\ \left|Y_m^{(f_c)}\right| d_K^{(k)}\left(Y_m^{(f_c)}\right)\Bigg\} \\
&= \frac{1}{N^2}\sum_{n=0}^{N-1}\mathrm{E}\left\{\left|Y_n^{(f_c)}\right|^2\ d_K^{(k)}\left(Y_n^{(f_c)}\right)\right\} + \frac{1}{N^2}\underbrace{\sum_{n=0}^{N-1}\sum_{m=0}^{N-1}}_{m\neq n}\mathrm{E}\left\{\left|Y_n^{(f_c)}\right| d_K^{(k)}\left(Y_n^{(f_c)}\right)\left|Y_m^{(f_c)}\right| d_K^{(k)}\left(Y_m^{(f_c)}\right)\right\} \\
&= \frac{1}{N^2}\sum_{n=0}^{N-1}\left(\mathrm{E}\left\{\left|Y_n^{(f_c)}\right|^2 d_K^{(k)}\left(Y_n^{(f_c)}\right)\right\} + \frac{1}{N^2}\underbrace{\sum_{n=0}^{N-1}\sum_{m=0}^{N-1}}_{m\neq n}\left(f^{(\mathcal{A},0,P)}\left(\psi_k+\phi_n^{(f_c)}\right)\ f^{(\mathcal{A},0,P)}\left(\psi_k+\phi_m^{(f_c)}\right)\right)\right. \\
&= \left(f^{(\mathcal{A},0,P)}\left(\psi_k\right)\right)^2 + \frac{1}{N^2}\sum_{n=0}^{N-1}\left[f^{(\mathcal{A},0,2)}\left(\psi_k+\phi_n^{(f_c)}\right) - f^{(\mathcal{A},0,P)}\left(\psi_k+\phi_n^{(f_c)}\right)^2\right]
\end{aligned} \tag{H.2}$$

where we have adopted the following compact notations:

$$f^{(\mathcal{A},0,2)}\left(\psi_k+\phi_n^{(f_c)}\right) \stackrel{\text{def}}{=} \mathrm{E}\Big\{\left|Y_n^{(f_c)}\right|^2 d_K^{(k)}\left(Y_n^{(f_c)}\right)\Big\}$$

and

$$\phi_n^{(f_c)} \stackrel{\text{def}}{=} 8\pi(f_c - f_0)n$$

Using (H.2), it follows that:

$$m_{f_c}^{(2)}[k] \stackrel{\text{def}}{=} \mathrm{E}\left\{\epsilon_k (f_c - f_0)^2\right\} = \frac{1}{N^2}\sum_{n=0}^{N-1}\left[f^{(\mathcal{A},0,2)}\left(\psi_k + \phi_n^{(f_c)}\right) - f^{(\mathcal{A},0,P)}\left(\psi_k + \phi_n^{(f_c)}\right)^2\right]$$

As far as the second order moment $\mathrm{E}\left\{a^{(\mathcal{A},(f_{c_1}-f_0),0)}\left(\psi_{k_1}\right)\cdot a^{(\mathcal{A},(f_{c_2}-f_0),0)}\left(\psi_{k_2}\right)\right\}$ is concerned, using the triangular pulse defined as follows:

$$\mathrm{tri}_{\frac{2\pi}{K}}(\cdot) = \begin{cases} 1 - |(\cdot)| & \text{for} \quad |(\cdot)| \le 2\pi/K \\ 0 & \text{for} \quad |(\cdot)| \ge 2\pi/K. \end{cases}$$

we have

$$\begin{aligned}
&\mathrm{E}\left\{a^{(\mathcal{A},(f_{c_1}-f_0),0)}\left(\psi_{k_1}\right)\cdot a^{(\mathcal{A},(f_{c_2}-f_0),0)}\left(\psi_{k_2}\right)\right\} \\
&\quad = \frac{1}{N^2}\cdot \mathrm{E}\left\{\left|Y_n^{(f_{c_1})}\right|\ \left|Y_n^{(f_{c_2})}\right|\ d_K^{(k_1)}\left(Y_n^{(f_{c_1})}\right) d_K^{(k_2)}\left(Y_n^{(f_{c_2})}\right)\right\} \\
&\qquad\quad + \frac{1}{N^2}\underbrace{\sum_{n=0}^{N-1}\sum_{m=0}^{N-1}}_{m\neq n}\left[\mathrm{E}\left\{\left|Y_n^{(f_{c_1})}\right|\ d_K^{(k_1)}\left(Y_n^{(f_{c_1})}\right)\right\}\mathrm{E}\left\{\left|Y_m^{(f_{c_2})}\right|\ d_K^{(k_2)}\left(Y_m^{(f_{c_2})}\right)\right\}\right] \\
&\quad = f^{(\mathcal{A},0,P)}\left(\psi_{k_1}\right)\ f^{(\mathcal{A},0,P)}\left(\psi_{k_2}\right) \\
&\qquad\quad + \frac{1}{N^2}\sum_{n=0}^{N-1}\left[f^{(\mathcal{A},0,2)}\left(\psi_{k_1} + \phi_n^{(f_{c_1})}\right)\ \mathrm{tri}_{\frac{2\pi}{K}}\left(\psi_{k_1} - \psi_{k_2} + \phi_n^{(f_{c_1})} - \phi_n^{(f_{c_2})}\right)\right. \\
&\qquad\quad \left. - f^{(\mathcal{A},0,P)}\left(\psi_{k_1} + \phi_n^{(f_{c_1})}\right)\ f^{(\mathcal{A},0,P)}\left(\psi_{k_2} + \phi_n^{(f_{c_2})}\right)\right]
\end{aligned} \tag{H.3}$$

where the last equality exploits the relation:

$$\begin{aligned}
&\mathrm{E}\left\{\left|Y_n^{(f_{c_1})}\right|\ \left|Y_n^{(f_{c_2})}\right|\ d_K^{(k_1)}\left(Y_n^{(f_{c_1})}\right) d_K^{(k_2)}\left(Y_n^{(f_{c_2})}\right)\right\} \\
&\qquad = \mathrm{E}\left\{\left|Y_n^{(f_{c_1})}\right|^2\ d_K^{(k_1)}\left(Y_n^{(f_{c_1})}\right)\right\}\cdot \mathrm{tri}_{\frac{2\pi}{K}}\left(\psi_{k_1} - \psi_{k_2} + \phi_n^{(f_{c_1})} - \phi_n^{(f_{c_2})}\right)
\end{aligned}$$

From (H.2) and (H.3) we have:

$$m_{f_c}^{(2)}[k_1,k_2] = \frac{1}{N^2}\sum_{n=0}^{N-1}\left[f^{(\mathcal{A},0,2)}\left(\psi_{k_1} + \phi_n^{(f_c)}\right)\delta_{k_1k_2} - f^{(\mathcal{A},0,P)}\left(\psi_{k_1} + \phi_n^{(f_c)}\right)\ f^{(\mathcal{A},0,P)}\left(\psi_{k_2} + \phi_n^{(f_c)}\right)\right]$$

where $\delta_{k_1 k_2}$ is the Kronecker delta, *i.e.* $\delta_{k_1 k_2} = 1$ if $k_1 = k_2$, otherwise $\delta_{k_1 k_2} = 0$. Furthermore:

$$m^{(1,1)}_{f_{c_1}, f_{c_2}}[k_1, k_2] = \frac{1}{N^2} \sum_{n=0}^{N-1} \Big[f^{(\mathcal{A},0,2)} \left(\psi_{k_1} + \phi_n^{(f_{c_1})} \right) \ \mathrm{tri}_{\frac{2\pi}{K}} \left(\psi_{k_1} - \psi_{k_2} + \phi_n^{(f_{c_1})} - \phi_n^{(f_{c_2})} \right) - f^{(\mathcal{A},0,P)} \left(\psi_{k_1} + \phi_n^{(f_{c_1})} \right) \ f^{(\mathcal{A},0,P)} \left(\psi_{k_2} + \phi_n^{(f_{c_2})} \right) \Big]$$

I A fast rough frequency offset estimator

The expected value of the accumulation function (4.10) when frequency compensation has not been applied, *i.e.* $f_c=0$, takes the following form:

$$\mathrm{E}\left\{a^{(\mathcal{A},(f_c-f_0),0)}\left(\psi_k\right)\right\}\Big|_{f_c=0} = \frac{1}{N}\sum_{n=0}^{N-1} f^{(\mathcal{A},\theta,P)}\left(\psi_k - 8\pi f_0 n\right) \tag{I.1}$$

From (I.1) we see that the accumulation function is a temporal average of N versions of the CPS, each one shifted by $8\pi f_0$ with respect to the preceding. Without loss of generality we can set $\theta=0$; by taking the DFT of both members we can write:

$$\mathrm{E}\left\{A^{(\mathcal{A},-f_0,0)}\left(m\right)\right\} = F^{(\mathcal{A},0)}\left(m\right)\cdot\frac{1}{N}\sum_{n=0}^{N-1}\exp(j8\pi f_0 mn)$$

Hence, after having approximated $F^{(\mathcal{A},0)}\left(m\right)\simeq \underset{\Phi}{G}^{(\mathcal{A},0)}\left(m\right)$, we can recast the frequency offset estimation as a nonlinear fitting problem between the function

$$\begin{aligned} U(mf_0) &\stackrel{\text{def}}{=} \frac{1}{N}\sum_{n=0}^{N-1}\exp(j8\pi f_0 mn) \\ &= \frac{1}{N}e^{-j8\pi f_0 m(N-1)/2}\,\frac{\sin(8\pi f_0 mN/2)}{\sin(8\pi f_0 m/2)} \end{aligned} \tag{I.2}$$

and the observations given by

$$\overline{U}_m = \frac{A^{(\mathcal{A},-f_0,0)}\left(m\right)}{\underset{\Phi}{G}^{(\mathcal{A},0)}\left(m\right)}$$

The rough estimate is thus obtained from the following minimization problem:

$$\hat{f}_0^{(c)} = \arg\min_{f_0}\sum_{m=0}^{L-1}\left(U(mf_0)-\overline{U}_m\right)^2$$

The nonlinear fitting can be applied on a limited number of points L, and here we have considered $L=8$. Numerical results have shown that the rough estimate $\hat{f}_0^{(c)}$ always remains within the interval $[f_0 \pm 10\Delta f_c]$, so that only a few number, typically 20, of frequency compensations are needed in the fine estimation stage.

References

[1] S. L. Hansen, "Large sample properties of the Generalized Methods of Moments", *Econometrica*, vol.50, pp.1029-1054, 1982.

[2] D. McFadden, Generalized Method of Moments, *Economics 240B Course Reader*, University of California, Berkeley (CA), USA, http://elsa.berkeley.edu/~mcfadden/e240b_sp03/e240b.html.

[3] Generalized Method of Moments Estimation, L. Matyas ed., Cambridge University Press, 1999.

[4] J. K. Tugnait, "Identification of Linear Stochastic Systems via Second- and Fourth-order Cumulant Matching", *IEEE Transactions on Information Theory*, vol.33, no.3, May 1987, pp.393-407.

[5] B. Porat, B. Friedlander, "Performance Analysis of Parameter Estimation Algorithms Based on High-order Moments", *Int. J. Adaptive Control Signal Processing*, vol.39, 1989, pp.191-229.

[6] B. Friedlander, B. Porat, "Asymptotically Optimal Estimation of MA and ARMA Parameter of Non-Gaussian Process from High-Order Moments", *IEEE Transactions on Automatic Control*, vol. 35, no. 1. January 1990, pp.27-35.

[7] Y. Rosen, B. Porat, "Optimal ARMA Parameter Estimation Based on the Sample Covariances for Data with Missing Observations", *IEEE Transactions on Information Theory*, vol. 35, no. 2. March 1989, pp.342-349.

[8] M. Ghogho, A Swami, A. K. Nandi, "Nonlinear least squares estimation for

harmonics in multiplicative and additive noise", *Elsevier Signal Processing*, vol. 78, 1999, pp.43-60.

[9] Digital Processing of Random Signals, B. Porat, Prentice Hall Information and System Sciences Series, 1994

[10] G.C Carter, "Coherence and time delay estimation", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol.29, June 1981, pp.463470.

[11] Directional Statistics, K. V. Mardia, P. E. Jupp, Wiley Series in Probability and Statistics, 1999

[12] F. Rice, B. Cowley, M. Rice, "Cramér-Rao lower bounds for QAM phase and frequency estimation", *IEEE Transactions on Communications*, vol.48, Sept. 2001.

[13] Linear statistical snference and its application, C. R. Rao, Wiley Series in Probability and Statistics, 1973

[14] L. Cohen, "The Scale Representation", *IEEE Transaction on Signal Processing*, Vol.41, no.12, December 1993

[15] G. Iacovitti, G. Scarano, "Discrete-time techniques for time delay estimation", *IEEE Transactions on Signal Processing*, vol.2, Feb. 1993.

[16] Yu T. Su, Ju-Ya Chen, "Carrier-to-Interference ratio measurement using moments or histograms", *IEEE Transactions on Communication*, vol.48, no.8, pp.1338-1346, August 2000.

[17] P. Gao, C. Tepedelenlioglu, "Practical issues in estimation over multiaccess fading channels with TBMA wireless sensor networks", *IEEE Transactions on Signal Processing*, vol.56, no.3, pp.1217-1229, March 2008.

[18] R.S. Luis, A. Teixeira, P. Monteiro, "Optical signal-to-noise ratio estimation using reference asynchronous histograms", *IEEE Journal of Lightwave Technology*, vol.56, no.3, pp.731-743, March 2008.

[19] S. Xiang, H. J. Kim and J. Huang, Senior Member, IEEE, "Invariant image watermarking based on statistical features in the low-frequency domain",

IEEE Transactions on Circuits and Systems for Video Technology, vol.18, no.6, pp.777-790, June 2008.

[20] S. Xiang, H. J. Kim and J. Huang, Senior Member, IEEE, "On scene segmentation and histograms-based curve evolution", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.31, no.9, pp.1708-1714, September 2009.

[21] Y. Wang, E. Serpedin, P. Ciblat, "Optimal Blind Nonlinear Least-Squares Carrier Phase and Frequency Offset Estimator for General QAM Modulation", *IEEE Transactions on Wireless Communications*, vol.2, September 2003.

[22] **Matrix computations** (3rd ed.), G.H. Golub and C. F. Van Loan, Johns Hopkins University Press, 1996.

[23] K.V. Cartwright, "Blind phase recovery in general QAM communication systems using alternative higher order statistics", *IEEE Signal Processing Letters*, vol.6, Dec. 1999.

[24] L. Chen, H. Kusaka, and M. Kominami, "Blind phase recovery in QAM communication systems using higher order statistics", *IEEE Signal Processing Letters*, vol.3, May 1996.

[25] M. Moeneclaey and G. de Jonghe, "ML-oriented NDA carrier synchronization for general rotationally symmetric signal constellations", *IEEE Transactions on Communications*, vol.42, Aug. 1994.

[26] 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, Aug. 2001.

[27] K.V. Cartwright, "Blind phase recovery in cross QAM communication systems with eighth-order statistics", *IEEE Signal Processing Letters*, vol.8, Dec. 2001.

[28] M. Morelli, U. Mengali, A.N.'D'Andrea, "Feed-forward estimation techniques for carrier recovery in 16-QAM modulation, in *Broadband Wireless*

Communications, M. Luise and S. Pupolin eds., New York: Springer-Verlag 1988.

[29] R. Lopez-Valcarce, C. Mosquera, "Blind carrier phase estimation for non-equiprobable constellations", *IEEE International Workshop on Signal Processing Advances in Wireless Communications, SPAWC 2006*, July 2006, Cannes, France.

[30] C.N. Georghiades, "Blind carrier phase acquisition for QAM constellations", *IEEE Transactions on Communications*, vol.45, no.11, Nov.1997.

[31] K.V. Cartwright, E.J. Kaminsky, "Blind phase recovery in cross QAM communication systems with the reduced-constellation eighth-order estimator (RCEOE)", *IEEE Global Telecommunications Conference, GLOBECOM 2005*, Nov. 2005, St.Luis, Missouri.

[32] G. Panci, P. Campisi, S. Colonnese, G. Scarano, "Blind phase recovery for QAM communications systems", *IEEE Transactions on Signal Processing*, April 2005.

[33] G. Panci, S. Colonnese, G. Scarano, "Asymptotically efficient phase recovery for QAM communication systems", ICASSP-2006, Toulouse, France, May 14-19, 2006.

[34] ETS 300 429, "Digital broadcasting systems for television, sound and data service; Framing structure channel coding and modulation for cable systems", Dec. 1994.

[35] C.W. Therrien, "Discrete Random Signals and Statistical Signal Processing", Prentice-Hall, 1992.

[36] P. Gao, C. Tepedelenlioglu, "SNR estimation for nonconstant modulus constellations", *IEEE Transactions on Signal Processing*, vol.53, Mar. 2005.

[37] C.N. Georghiades, W.S. Yuan, "Rapid carrier acquisition from baud-rate samples", *IEEE Transactions on Communications*, vol.47, April 1999.

[38] P. Ciblat, P. Loubaton, E. Serpedin, G.B. Giannakis, "Performance analysis of blind carrier frequency offset estimators for noncircular transmissions

through frequency-selective channels", *IEEE Transactions on Signal Processing*, vol.50, no.1, Jan. 2002.

[39] P. Ciblat, L. Vandendorpe, "Blind carrier frequency offset estimation for noncircular constellation-based transmissions", *IEEE Transactions on Signal Processing*, vol.51, no.5, May 2003.

[40] P. Ciblat, E. Serpedin, Y. Wang, "On a blind fractionally sampling-based carrier frequency offset estimator for noncircular transmissions", *IEEE Signal Processing Letters*, vol.10, April 2003.

[41] G. Panci, S. Colonnese, S. Rinauro, G. Scarano, "Gain-control-free near efficient phase acquisition for QAM constellations", *IEEE Transactions on Signal Processing*, vol.56, no.7, July 2008.

[42] S. Colonnese, G. Panci, S. Rinauro, G. Scarano, "Asymptotically efficient phase recovery for QAM communication systems", *Int. Symp. Wireless Communication Systems - ISWCS 2007*, Trondheim, (Norway), 2007.

[43] S. Colonnese, G. Panci, S. Rinauro, G. Scarano, "High SNR Performance Analysis of a Blind Frequency Offset Estimator for Cross QAM Communication", *ICASSP 2008*, Las Vegas (USA) 2008.

[44] R.A. Wiggins, "Minimum entropy deconvolution", *Geoexploration*, Vol. 16, 1978.

[45] S. Bellini, "Bussgang techniques for blind deconvolution and restoration," in Blind Deconvolution, S. Haykin ed., Prentice-Hall, 1994.

[46] M. Kendall, A. Stuart, J. K. Ord, S. Arnold, *Kendall's Advanced Theory of Statistics: Volume 2A – Classical Inference and the Linear Model*, vol.2, Feb. 1993.

[47] Fundamentals Of Statistical Signal Processing-Estimation Theory, S M. Kay, Prentice Hall Signal Processing Series, 1993

[48] A. N. D'Andrea, U. Mengali, R. Reggianini, "The modified Cramer Rao bound and its applications to synchronization problems", *IEEE Transactions on Communications*, vol.42, Feb./Mar./Apr. 1994.

[49] Monte Carlo Statistical Methods, C. P. Robert and G. Casella, Springer Verlag, New York, 2001.

List of Figures