SAMPLE COVARIANCE MATRIX PARAMETER ESTIMATION: CARRIER FREQUENCY, A CASE STUDY

Javier Villares, Gregori Vázquez

Department of Signal Theory and Communications, Universitat Politècnica de Catalunya
c/ Jordi Girona 1-3, Mòdul D-5, Campus Nord UPC, 08034 Barcelona (Spain)
e-mail:{javi,gregori}@gps.tsc.upc.es

ABSTRACT

This paper presents a novel approach to parameter estimation based on the sample covariance matrix linear processing. The originality of this framework relies upon two facts: firstly, the gaussian assumption about the nuisance parameters is avoided and, secondly, quadratic feed-forward schemes are designed saving the complexity and delay of conventional ML-based algorithms that carry out an exhaustive search throughout the whole parameter range. This second aim is achieved adopting a Bayesian perspective in which the parameter of interest is modeled as a random variable of known a priori distribution. The Bayesian approach allows us to establish certain optimality criteria (mean squared error, bias and variance) yielding to estimation schemes with the best performance on the average, that is, with respect to the assumed prior of the parameter.

In order to illustrate the proposed theory, we address the problem of frequency estimation in digital communications. This example has been chosen because its formulation encompasses several problems of special interest such as non-data-aided open-loop carrier synchronization, direction-of-arrival estimation in narrow-band uniform linear arrays and, if the signal is processed in the frequency domain, timing recovery and time-of-arrival estimation in positioning systems, as well.

1. INTRODUCTION

It is well-known that in low-SNR scenarios, the stochastic (or unconditional) Maximum Likelihood estimator is *locally* (nearby a reference value of the unknown parameter) quadratic on the observed vector $\mathbf{y}$ [1][2]. Thus, the ML theory supplies *large sample* efficient second-order trackers for the steady-state. However, in general, the optimal estimation scheme is still unknown out of these assumptions. In particular, when we deal with low-complexity implementations, efficient estimators may not even exist below a given SNR threshold due to the so-called *outliers* [3]. Under these circumstances, the utilisation of any a priori knowledge about the unknown parameter can be useful to atenuate the outliers effect [4]. From the Bayesian estimation theory, it is kwnon that the Minimum Mean Squared Error (MMSE) estimator is given by the a posteriori mean of the parameter, conditioned on the observed data [5]. Unfortunately, an analytical expression cannot be normally obtained. Although other non-linearities could result in better performance, we

will focus on those estimators processing linearly the sample covariance matrix $\widehat{\mathbf{R}} \doteq \mathbf{y}\mathbf{y}^H$, which concentrates the instantaneous second-order statistics of the observed vector $\mathbf{y}$. The reasoning for the quadratic constraint comes from the fact that it is the *smallest* (and hence suited for low-SNRs), *affordable* non-linearity yielding to a parameter identifiable problem in spite of the random nuisance parameters [2].

2. SIGNAL MODEL

The received complex envelope in a single carrier per channel (SCPC) system can be represented as follows:

$$y(t) = e^{j(\phi + \omega t)} \sum_{k=-\infty}^{\infty} c_k g(t - kT) + w(t) \qquad (1)$$

where $\nu \doteq \omega/2\pi T$ is the carrier frequency error we aim to estimate (normalized to the symbol period T), ϕ is an arbitrary random variable setting the phase origin at time $t = \omega/\phi$, $\{c_k\}$ is the sequence of zero-mean uncorrelated symbols, $g(t)$ the shaping pulse and, $w(t)$ the additive gaussian noise term, possibly colored due to previous filtering.

Firstly, the received signal is sampled taking samples each $T_s \doteq T/N_{ss}$ seconds with N_{ss} the integer oversampling factor. The signal timing is assumed to be perfectly determined. At a given time (say $t=0$, for instance) N consecutive samples are delivered to the estimator. The vector collecting these samples $\mathbf{y} \doteq [y(0), \dots, y((N-1)T_s)]^T$ can be expressed as:

$$\mathbf{y} = \mathbf{A}(\nu)\mathbf{x} + \mathbf{w} \qquad (2)$$

where $\mathbf{x}$ gathers those symbols having contribution in the observed interval (absorbing the phase ambiguity ϕ) and $\mathbf{A}(\nu)$ is the transfer matrix containing the fraction of the shaping pulses conveying these symbols.

Finally, the parameter of interest ν is modeled as a uniform random variable in the interval $(-\Delta/2, \Delta/2)$ where the uncertainty range $\Delta \leq N_{ss}$ (Nyquist bandwidth) is all the *a priori* knowledge the designer has about the value of ν.

3. FREQUENCY-OFFSET QUADRATIC ESTIMATION

According to the introductory discussion, the generic expression of any second-order estimator is given by:

$$\widehat{\nu} = B + Tr\left\{\mathbf{M}\widehat{\mathbf{R}}\right\} = B + \mathbf{m}^H \widehat{\mathbf{r}} \qquad (3)$$

where B and $\mathbf{M}$ are the estimator coefficients the designer should select under certain optimality criterion. Note that

the linear term is omitted due to the symbols zero mean. To facilitate the estimators deduction, in (3) we have vectorized $\mathbf{M}$ and $\widehat{\mathbf{R}}$ so that $\mathbf{m} \doteq vec\{\mathbf{M}^H\}$ and $\widehat{\mathbf{r}} \doteq vec\{\widehat{\mathbf{R}}\}$ where $(\cdot)^H$ stands for the transpose conjugate.

First of all, the estimator mean value is a function of the signal covariance matrix $\mathbf{R}(\nu)$ as indicated below:

$$E\{\widehat{\nu}\} = B + Tr\{\mathbf{M}\mathbf{R}(\nu)\} = B + \mathbf{m}^H \mathbf{r}(\nu) \qquad (4)$$
$$\mathbf{R}(\nu) = \mathbf{A}(\nu)\mathbf{A}^H(\nu) + \mathbf{R}_w$$

where $\mathbf{r}(\nu) \doteq vec\{\mathbf{R}(\nu)\}$ is the vectorization of $\mathbf{R}(\nu)$ and $\mathbf{R}_w \doteq E\{\mathbf{w}\mathbf{w}^H\}$ the noise covariance matrix.

Having in mind (4), next we formulate the estimator bias, variance and MSE that will be used hereafter to design the different estimation strategies and assess their performance:

$$BIAS^2(\nu) \doteq \|E\{\widehat{\nu}\} - \nu\|^2 = \left\|B + \mathbf{m}^H \mathbf{r}(\nu) - \nu\right\|^2$$
$$VAR(\nu) \doteq E\|\widehat{\nu} - E\{\widehat{\nu}\}\|^2 = E\left\|\mathbf{m}^H(\widehat{\mathbf{r}} - \mathbf{r}(\nu))\right\|^2$$
$$MSE(\nu) \doteq BIAS^2(\nu) + VAR(\nu) \qquad (5)$$

Regarding the above equations, we see that the same value of B minimizes both the bias and the MSE, as the estimator variance is independent of B. Concretely, if we choose B in order to minimize the overall bias, $BIAS^2 \doteq E_\nu\{BIAS^2(\nu)\}$, that is, the bias across the whole frequency error range (weighted by the known prior), we obtain that:

$$\frac{\partial}{\partial B^*} BIAS^2 = B + E_\nu\{\mathbf{m}^H \mathbf{r}(\nu) - \nu\} = B + \mathbf{m}^H \mathbf{r} = 0 \quad (6)$$

where $E_\nu\{\nu\} = 0$, $\mathbf{r} \doteq E_\nu\{\mathbf{r}(\nu)\}$ and, thus, (3) becomes:

$$\widehat{\nu} = \mathbf{m}^H(\widehat{\mathbf{r}} - \mathbf{r}) \qquad (7)$$

If we operate now the expressions in (5), we obtain that:

$$BIAS^2(\nu) = \mathbf{m}^H \mathbf{Q}_{bias}(\nu)\mathbf{m} - 2\operatorname{Re}\left\{\mathbf{m}^H \mathbf{s}(\nu)\right\} + \|\nu\|^2$$
$$VAR(\nu) = \mathbf{m}^H \mathbf{Q}_{var}(\nu)\mathbf{m} \qquad (8)$$

with the following definitions:

$$\mathbf{Q}_{bias}(\nu) \doteq (\mathbf{r}(\nu) - \mathbf{r})(\mathbf{r}(\nu) - \mathbf{r})^H$$
$$\mathbf{Q}_{var}(\nu) \doteq E\left\{(\widehat{\mathbf{r}} - \mathbf{r}(\nu))(\widehat{\mathbf{r}} - \mathbf{r}(\nu))^H\right\} = \mathbf{R}^T(\nu) \otimes \mathbf{R}(\nu) +$$
$$+ \mathbf{B}(\nu)\mathbf{K}\mathbf{B}^H(\nu)$$
$$\mathbf{s}(\nu) \doteq \mathbf{r}(\nu)\nu \qquad (9)$$

where $\mathbf{B}(\nu) \doteq \mathbf{A}^*(\nu) \otimes \mathbf{A}(\nu)$ and we have introduced the modulation fourth-order cumulant matrix $\mathbf{K}$ given by:

$$\mathbf{K} = E\left\{vec\{\mathbf{X}\}vec^H\{\mathbf{X}\}\right\} - vec\{\mathbf{I}\}vec^H\{\mathbf{I}\} - \mathbf{I} \quad (10)$$

where $\mathbf{X} \doteq \mathbf{x}\mathbf{x}^H$ and $\mathbf{I}$ denotes the identity matrix. It is well-known that $\mathbf{K}$ would vanish if the nuisance parameters were normally distributed. However, this does not happen, for instance, in digital communications, and matrix $\mathbf{K}$ provides the complete non-gaussian information about the discrete symbols that second-order estimators (3) are able to exploit. In the case of linear modulations, such as *PSK*, *QAM* or, in general, *APSK*, matrix $\mathbf{K}$ reduces to:

$$\mathbf{K} = (\rho - 2)diag\{vec\{\mathbf{I}\}\} \qquad (11)$$

where the $\rho \doteq E\{|x_i|^4\}/E^2\{|x_i|^2\}$ ($\forall i$) is the fourth- to second-order moment ratio (specific of the modulation under consideration), and $diag(\cdot)$ converts a vector into a diagonal matrix.

In order to determine the estimators coefficients in $\mathbf{m}$, we will adopt a Bayesian approach in which the parameter is a random variable we average with the purpose of deducing schemes that work properly on average and exploiting the available statistical knowledge about the unknown parameter (prior). The Bayesian counterparts of the performance indicators in (5) are given next:

$$BIAS^2 = \mathbf{m}^H \mathbf{Q}_{bias}\mathbf{m} - 2\operatorname{Re}\left\{\mathbf{m}^H \mathbf{s}\right\} + \sigma_\nu^2 \qquad (12)$$
$$VAR = \mathbf{m}^H \mathbf{Q}_{var}\mathbf{m}$$
$$MSE = BIAS^2 + VAR = \mathbf{m}^H \mathbf{Q}_{mse}\mathbf{m} - 2\operatorname{Re}\left\{\mathbf{m}^H \mathbf{s}\right\} + \sigma_\nu^2$$

with

$$\mathbf{Q}_{bias} \doteq E_\nu\left\{\mathbf{r}(\nu)\mathbf{r}^H(\nu)\right\} - \mathbf{r}\mathbf{r}^H$$
$$\mathbf{Q}_{var} \doteq E_\nu\{\mathbf{Q}_{var}(\nu)\} = E_\nu\left\{\mathbf{R}^T(\nu) \otimes \mathbf{R}(\nu)\right\} +$$
$$+ E_\nu\left\{\mathbf{B}(\nu)\mathbf{K}\mathbf{B}^H(\nu)\right\}$$
$$\mathbf{Q}_{mse} \doteq \mathbf{Q}_{bias} + \mathbf{Q}_{var} \qquad (13)$$
$$\mathbf{s} \doteq E_\nu\{\mathbf{r}(\nu)\nu\}$$
$$\sigma_\nu^2 \doteq E_\nu\{\nu^2\} = \frac{1}{3}\Delta^2 \qquad (14)$$

Up to now, the formulation is totally general and will encompass any estimation problem following the linear model stated in section 2. In the case of frequency estimation, the above matrices ($\mathbf{Q}_{bias}$, $\mathbf{Q}_{var}$ and $\mathbf{s}$) can be calculated analytically because of the parameter phasor dependence and the following equality:

$$\mathbf{G}(\nu) \doteq \mathbf{A}(\nu)\mathbf{A}^H(\nu) = \mathbf{E}(\nu) \odot \mathbf{G}(0) \qquad (15)$$

where $\odot$ stands for the element-wise Hadamard product and $\mathbf{E}(\nu)$ is defined as:

$$[\mathbf{E}(\nu)]_{i,j} = e^{j2\pi\nu(i-j)/N_{ss}} \qquad (16)$$

Having in mind this result, we have that:

$$\mathbf{Q}_{bias} = \left(\mathbf{E}_q - \mathbf{e}\mathbf{e}^H\right) \odot \mathbf{g}(0)\mathbf{g}^H(0)$$
$$\mathbf{Q}_{var} = \mathbf{R}^T \otimes \mathbf{R} - \left(\mathbf{E}^T \otimes \mathbf{E}\right) \odot \left(\mathbf{G}^T(0) \otimes \mathbf{G}(0)\right) +$$
$$+ \mathbf{E}_q \odot \mathbf{B}(0)\mathbf{K}\mathbf{B}^H(0)$$
$$\mathbf{s} = \mathbf{e}_s \odot \mathbf{g} \qquad (17)$$

where we have introduced matrices $\mathbf{E} \doteq E_\nu\{\mathbf{E}(\nu)\}$, $\mathbf{E}_s \doteq E_\nu\{\mathbf{E}(\nu)\nu\}$, $\mathbf{E}_q \doteq E_\nu\{\mathbf{e}(\nu)\mathbf{e}^H(\nu)\}$ and its vectorized versions $\mathbf{e} \doteq vec\{\mathbf{E}\}$, $\mathbf{e}_s \doteq vec\{\mathbf{E}_s\}$. From (16), analytical expressions for these matrices are straighforward.

4. BIAS MINIMIZATION

In this section we examine the control the designer has over the estimator bias. Because of the non-linear relationship between the parameter and the observed data (16), in general,
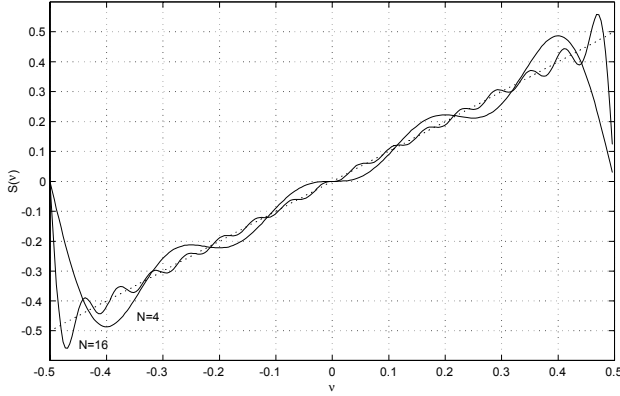
Figure 1: Frequency estimator mean response $S(\nu)$ as a function of N for $\Delta = N_{ss} = 1$

the overall bias in (12) cannot be cancelled out [6]. Then, this section is devoted to minimizing this bias and evaluating its magnitude in the case of frequency estimation.

The estimator coefficients $\mathbf{m}$ minimizing (12) are readily obtained from the derivative of $BIAS^2$, concluding that the equation any estimator $\mathbf{m}$ must hold to deliver mimimum-biased estimates is:

$$\mathbf{Q}_{bias}\mathbf{m} = \mathbf{s} \qquad (18)$$

that can be rewritten, after some trivial manipulations, as follows:

$$E_\nu \{\mathbf{G}(\nu) S(\nu)\} = E_\nu \{\mathbf{G}(\nu)\nu\} \qquad (19)$$

where $S(\nu) = \mathbf{m}^H (\mathbf{r}(\nu) - \mathbf{r})$ denotes the estimator mean value (4) as a function of the parameter ν. If we look at (19), it is clear that an unbiased estimator will comply with (18). Unfortunately, this is not mostly possible and (18) supplies the least squares fitting of $S(\nu)$ to the ideal linear response $S(\nu) = \nu$ within the prior domain (i.e., $|\nu| < \Delta/2$).

Furthermore, if some elements in $\mathbf{G}(\nu)$ are connected by an affine transformation, i.e., $[\mathbf{G}(\nu)]_{i_2,j_2} = C_a [\mathbf{G}(\nu)]_{i_1,j_1} + C_b$ for any value of C_a and C_b, the system of equations in (19) becomes underdetermined and hence $\mathbf{Q}_{bias}$ is rank-deficient. Indeed, this is exactly what happens in the frequency estimation case since the diagonals of $\mathbf{G}(\nu)$ share the same phasor (16). Thus, it is possible to reduce (19) to $2N - 1$ equations, one per diagonal of $\mathbf{G}(\nu)$, as indicated next:

$$E_\nu \left\{ S(\nu) e^{j2\pi\nu n/N_{ss}} \right\} = E_\nu \left\{ \nu e^{j2\pi\nu n/N_{ss}} \right\}$$

$$\int_{-R/2}^{R/2} V(f) e^{j2\pi fn} df = N_{ss} \int_{-R/2}^{R/2} f e^{j2\pi fn} df \qquad (20)$$

where $n \in (-N, N)$, $f \doteq v/N_{ss}$, $R \doteq \Delta/N_{ss}$ is the carrier uncertainty relative to the Nyquist bandwidth N_{ss} and, $V(f) \doteq Tr\{\mathbf{M}\mathbf{G}(\nu)\}$ is the Fourier transform of the sequence $v(n)$ defined next:

$$v(n) \doteq FT^{-1}\{V(f)\} = \sum_i [\mathbf{M}]_{i,i+n} [\mathbf{G}(0)]_{i,i+n} \qquad (21)$$

Notice that in (20) we have taken into account that $S(\nu) = V(f) - C$ where C must be null to guarantee the odd symmetry of the harmonic expansion of f in the right-hand side of (20).

Thus, equation (20) states that the $2N - 1$ central terms of the discrete Fourier series of $N_{ss}f$ and $V(f)$, filtered in the interval $\pm R/2$, must be identical in order to minimize the estimator bias. Ideally, if N were arbitrarily long, (20) would imply the equalization of $S(\nu)$ and ν within the prior interval $|\nu| < \Delta/2$ or, in other words:

$$V(f) = \lim_{N \to \infty} \sum_{n=-N+1}^{N-1} v(n) e^{-j2\pi fn} = N_{ss}f \qquad (22)$$

for $|f| < R/2$, whatever the value of R. However, since N is finite, the value at which the above Fourier series can be truncated without noticiable distortion is a function of the ratio $R = \Delta/N_{ss}$; the smaller R, the less terms are required for the same distortion of $S(\nu)$. In the limit $(R \to 0)$, the Taylor expansion of (22) around $f = 0$ ensures that $N = 2$ is sufficient to hold exactly (22) with $v(1) = -v(-1) = iN_{ss}/(2\pi)$. Otherwise, if (22) is truncated taking too few elements, $S(\nu)$ will suffer from ripple and the Gibbs effect, i.e., the overshooting at the discontinuity points $(|\nu| = \pm\Delta/2)$, as shown in Fig. 1 for the most critical situation in which $R = 1$. The reader is referred to [6] for additional simulations on the estimator bias as a function of the parameter R.

Unfortunately, although we had an infinite observation $(N \to \infty)$, the non-zero signal bandwidth constitutes another limitation to the liniarization of $S(\nu)$. Returning to the definition of $v(n)$ in (21), we see that matrix $\mathbf{G}(0)$, whose diagonals are composed of the N_{ss} synchronous components of the shaping pulse autocorrelation (15), acts as a "temporal" window over the actual sequence $v(n)$. Therefore, the minimization of the bias is limited by the *effective* duration of this autocorrelation whatever the value of N. Moreover, as the minimum Nyquist bandwidth in communications is $1/T$ Hz, it follows that the main lobe of the signal autocorrelation lasts $2/T$ sec. and, thus, in practice the Fourier series in (22) becomes truncated approximately at $N = N_{ss}$.

5. MINIMUM MSE AND VAR ESTIMATORS

In the previous section the condition to minimize the term $BIAS^2$ in (12) was obtained. Depending on whether we impose this condition or not, two different solutions can be deduced. On the one hand, if bias is not acceptable, among the set of all the estimators holding the minimum bias constraint in (18), the one minimizing VAR, and hence $MSE = VAR + \min\{BIAS^2\}$, is obtained from (12) solving the following constrained optimization problem:

$$\mathbf{m}_{var} \doteq \arg\min_{\mathbf{m}} \left\{ VAR + (\mathbf{s} - \mathbf{Q}_{bias}\mathbf{m})^H \boldsymbol{\lambda} \right\} =$$

$$= \mathbf{Q}_{var}^{-1}\mathbf{Q}_{bias} \left(\mathbf{Q}_{bias}\mathbf{Q}_{var}^{-1}\mathbf{Q}_{bias}\right)^{\#} \mathbf{s} \doteq \mathbf{P}^H\mathbf{s} \qquad (23)$$

where the inversion of $\mathbf{Q}_{var}$ is guaranteed if the noise covariance matrix $\mathbf{R}_w$ is positive definite. The minimum-bias constraint (18) is imposed in (23) by the vector of Lagrange's multipliers $\boldsymbol{\lambda}$. Notice that the Moore-Penrose pseudo-inverse in (23) provides the minimum-norm solution to the underdetermined system of equations studied in section 4.

If (23) is plugged into (12), we have that the estimator minimum bias is given by:

$$\min\{BIAS^2\} = \sigma_\nu^2 - \mathbf{s}^H\mathbf{P}\mathbf{s} = \sigma_\nu^2 - \mathbf{s}^H\mathbf{Q}_{bias}^{\#}\mathbf{s} \qquad (24)$$
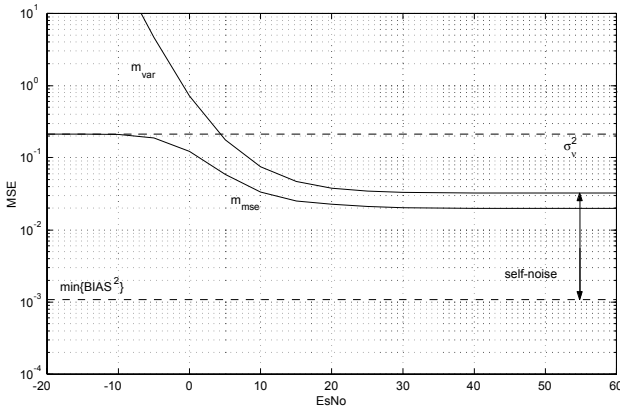
Figure 2: MSE for the two estimators deduced in section 5 for the Minimun Shift Keying (MSK) modulation with $N = 6$, $N_{ss} = 2$ and $\Delta = 0.8$

Figure 3: MSE as a function of the observation length (N) for MSK, $EsNo$=40dB, N_{ss}=2 and $\Delta = 0.8$

as the projection matrix $\mathbf{P}$ given by the constrained solution in (23) can be replaced by any other projector onto the subspace of minimum bias generated by matrix $\mathbf{Q}_{bias}$.

On the other hand, if the bias constraint is avoidable, the minimum MSE estimator and the associated MSE are:

$$\mathbf{m}_{mse} \doteq \arg\min_{\mathbf{m}} \{MSE\} = \mathbf{Q}_{mse}^{-1}\mathbf{s}$$
$$\min\{MSE\} = \sigma_\nu^2 - \mathbf{s}^H\mathbf{Q}_{mse}^{-1}\mathbf{s} \qquad (25)$$

The above estimator makes a trade-off between bias and variance so that it yields biased estimates with the aim of reducing the estimates variability in those noisy scenarios in which the variance contribution is dominant.

In figures 2 and 3 the two solutions are compared in terms of their MSE. In figure 2 we can observe how the second-order frequency estimators proposed herein suffer from self-noise at high SNRs due to the frequency-offset uncertainty. On the contrary, frequency error detectors in closed-loop schemes were found to be self-noise free [1]. The self-noise variability precludes the equivalence of both solutions at high SNRs and, thus, the estimator $\mathbf{m}_{mse}$ can outperform $\mathbf{m}_{var}$.

On the other hand, if the estimators performance is evaluated as a function of the observation length N (figure 3), we observe that, as the variability is averaged out ($\lim_{N\to\infty} VAR = 0$), both solutions converge and the prevalent, systematic error is the residual bias (section 4), whose minimum value is given by:

$$\lim_{N\to\infty} MSE = \sigma_\nu^2 - \lim_{N\to\infty} \mathbf{s}^H\mathbf{Q}_{bias}^{\#}\mathbf{s} \qquad (26)$$

Consequently, the estimators consistency in terms of MSE is solely guaranteed if $R = \Delta/N_{ss} \to 0$ (section 4).

6. CONCLUSIONS

In this paper the design and evaluation of quadratic estimators for the problem of frequency error estimation has been studied. We showed that quadratic unbiased estimators within the given parameter range do not exist unless the sampling rate is much higher than the maximum frequency error. Based on classical Fourier analysis, we found that the residual bias results from th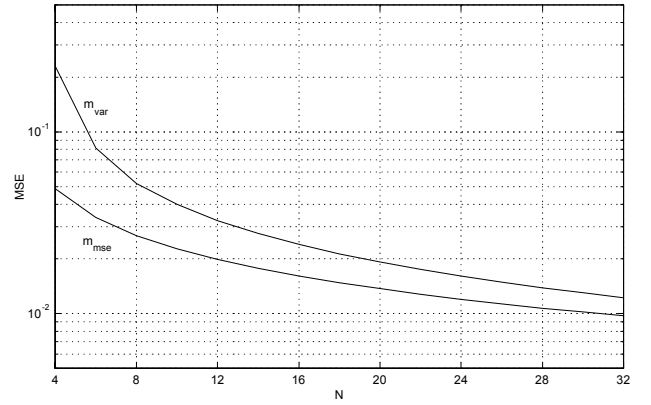e truncation of the harmonic expansion of the ideal, unbiased estimator mean response within the prior range. Likewise, the resulting distortion was related to the signal bandwidth and the aforementioned oversampling.

Depending on whether the minimum-bias constraint is imposed or not, two Bayesian estimators were deduced that minimize the average variance and MSE with respect to the available prior, respectively. Analytical expressions for both solutions and their performance were provided taking into account the actual statistics of the nuisance parameters. Simulations showed that the unconstrained solution outperforms its competitor as it makes a trade-off between bias and variance. Moreover, we observed that the performance of the minimum MSE estimator is bounded by the prior variance in low-SNR scenarios where the occurrence of outliers is likely. On the other hand, simulations also showed how second-order frequency estimators suffer from self-noise at high SNRs due to the uncertainty of the parameter. Finally, a large sample study for both solutions pointed out their asymptotical convergence.

7. REFERENCES

[1] G. Vázquez, J. Riba. *"Non-Data-Aided Digital Synchronization"*. In *Signal Processing Advances in Wireless Communications*, Vol. II, Chapter 9. Prentice-Hall, 2000.

[2] B. Ottersten, M. Viberg, P. Stoica. *"Exact and Large Sample Maximum Likelihood Techniques for Parameter Estimation and Detection"*. In *Radar Array Processing*, Chapter 4, pp. 99-151. Springer-Verlag, 1993.

[3] D.C. Rife, R.R. Boorstyn. *"Single Tone Parameter Estimation from Discrete-Time Observations"*. IEEE Trans. on Inf. Theory, Sep 1974.

[4] K. L. Bell, Y. Steinberg, Y. Ephraim, H.L. Van Trees. *"Extended Ziv-Zakai Lower Bounds for Vector Parameter Estimation"*. IEEE Trans. on Inf. Theory. Mar 1997.

[5] S. M. Kay. *"Fundamentals of Statistical Signal Processing. Estimation Theory"*. Prentice-Hall, 1993

[6] J. Villares, G. Vázquez. *"Sample Covariance Matrix Based Parameter Estimation for Digital Synchronization"*. IEEE Proc. of Globecom. Taipei (Taiwan). Nov 2002.