PARAMETER ESTIMATION OF BINARY CPM SIGNALS

Jaume Riba and Gregori Vazquez

Department of Signal Theory and Communications. Universitat Politècnica de Catalunya.
Campus Nord, Ed. D5, Jordi Girona 1 i 3, 08034, Barcelona (Spain). Email: jriba,gregori@gps.tsc.upc.es

ABSTRACT

Estimation of frequency and symbol timing in Continuous Phase Modulated (CPM) signals is investigated in this paper. Several well-known statistical approaches, classically applied to the sensor array problem, are used to derive Non-Data-Aided (NDA) algorithms under a unifying general framework (*estimation-directed*). A new cost function is proposed which is shown to provide a good compromise between additive and pattern noise cancellation, when the additive noise power is unknown.

1. INTRODUCTION

CPM schemes are attractive due to their high spectral efficiency and constant envelope. However, synchronization of these signals is a crucial issue (see [1] and references therein for an exhaustive review). The application of the maximum likelihood principle is mathematically difficult in the case of CPM signals. To circumvent this problem, some algorithms are derived through approximations or employing heuristic arguments.

Frequency and timing synchronization algorithms are typically categorized in Decision-Directed (DD), Data-Aided (DA) and Non-Data-Aided (NDA) methods. While DD and DA schemes offer better tracking and acquisition performance respectively, NDA methods are preferred when the decisions are not available or not reliable, and the data is not known. NDA algorithms offer the additional advantage of being phase-independent, thus avoiding spurious locks and prolonged acquisitions caused by complex iteractions between phase and frequency and/or phase and timing correction algorithms. Additionally, simple symbol-by-symbol decisions cannot be obtained from CPM waveforms, for which the DD schemes becomes more complicated.

In this paper we formulate the problem of synchronization of CPM signals in the same manner as several important problems in the field of signal processing, as the problem of direction finding with narrow-band sensor arrays. This leads naturally to the definition of a unifying general framework for the problem at hand, referred to as *Estimation-Directed* (ED) approach. As a case study, we focus on Gaussian Minimum Shift Keying (GMSK) modulation, which is the modulation adopted in the GSM European cellular mobile digital system. A recent work by the authors on this topic can be found in the book [2].

This work was supported by Spanish Goverment Grants TIC98-0703, TIC99-0849, TIC2000-1025 and FIT-070000-2000-649, and Generalitat de Catalunya Grant 2000SGR-00083.

2. BACKGROUND

Several important problems in the signal processing field can be reduced to estimating the parameters in the following model:

$$\mathbf{r} = \mathbf{A}(\lambda)\mathbf{x} + \mathbf{w} \tag{1}$$

where, λ is the parameter of interest, $\mathbf{A}(\lambda)$ is the transfer matrix dependent on λ and $\mathbf{w}$ is the white noise vector ($E\left[\mathbf{w}\mathbf{w}^H\right] = \sigma^2\mathbf{I}$, $\sigma^2 = 2N_0F_s$). As shown in [3], using the Laurent expansion (see [1] and references therein), the parameter estimation of binary CPM signals can be formulated in terms of the above model, where $\lambda = f$ (frequency) or $\lambda = \tau$ (timing), depending on the problem considered. The transfer matrix is given by:

$$\begin{aligned} \mathbf{A}(\lambda) &\doteq \begin{bmatrix} \mathbf{A}_o(\lambda), & \mathbf{A}_1(\lambda), & ..., & \mathbf{A}_{J-1}(\lambda) \end{bmatrix} \\ \mathbf{A}_m(\lambda) &\doteq [\mathbf{a}_{m,o}(\lambda), \mathbf{a}_{m,1}(\lambda), ..., \mathbf{a}_{m,L-1}(\lambda)] \end{aligned} \tag{2}$$

where the column vectors are:

$$\begin{aligned} \mathbf{a}_{m,i}(\lambda) \doteq\ & e^{j2\pi f(iT-T_o)} \\ & \big[g_m(-iT-\tau), g_m(-iT+T_s-\tau)e^{j2\pi fT_s}, ..., \\ & g_m(-iT+(M-1)T_s-\tau)e^{j2\pi fT_s(M-1)}\big]^T \end{aligned}$$

In this model, $g_m(t)$ ($m = 0, 1, ..., J-1$), where $J = 2^{L-1}$, are the so-called pseudo-pulses, whose number and shape is specific of the kind of modulation, and T_o is a constant that reflects the arbitrary time origin of the problem. The observation interval (the length of $\mathbf{r}$) is limited to M samples, being M arbitrarily fixed by the synchronizer designer[1]:

$$\mathbf{r}_t = \left[r(t), r(t+T_s), \ldots, r(t+(M-1)T_s)\right]^T \tag{3}$$

where $r(t)$ is the received low pass equivalent signal. The vector $\mathbf{x}$ of signals is:

$$\begin{aligned} \mathbf{x} &\doteq \begin{bmatrix} \mathbf{x}_0^T & \mathbf{x}_1^T & ... & \mathbf{x}_{J-1}^T \end{bmatrix}^T \\ \mathbf{x}_m &\doteq [x_{m,o}, x_{m,1}, ..., x_{m,L-1}]^T \\ x_{m,i} &\doteq c_{m,i}e^{j\theta_o} \end{aligned} \tag{4}$$

where $c_{m,n}$ are the pseudo-symbols, which verify that:

$$\mathbf{\Gamma} = E\left[\mathbf{x}\mathbf{x}^H\right] = \mathbf{I}. \tag{5}$$

[1]The dependence of $\mathbf{r}$ on t, will be explicited only when necessary.

and θ_o is the carrier phase error. We are assuming that the number of samples per symbol $N_{ss} = T/T_s$ is integer.

All the estimators presented in this paper are based on the sample covariance matrix of the received signal waveform, denoted by $\widehat{\mathbf{R}}$. As we are dealing with over sampled digital communications signals, which are cyclostationary, $\widehat{\mathbf{R}}$ is computed by means of a synchronized time average as follows:

$$\widehat{\mathbf{R}} = \sum_{k=0}^{N_s-1} \mathbf{r}_{kT}\mathbf{r}_{kT}^H \tag{6}$$

where N_s is the number of actual symbols of the signal from which to compute a single estimate of λ.

3. MAXIMUM LIKELIHOOD (ML) APPROACHES

3.1. Exact stochastic ML (ESML)

The ESML derivation of NDA algorithms involves two steps: i) computation of the joint ML function of the parameter of interest and the nuisance parameters, which include the pseudo-symbols; and ii) average of the joint ML function with respect to the (complete) statistics of the pseudo-symbols. Unfortunately the second step is difficult, and we are compelled to make approximations. A usual approximation is to assume that the SNR is low. The resulting algorithms are highly inefficient at moderate E_s/N_o, especially for non-linear modulation formats.

3.2. Gaussian stochastic ML (GSML)

If we assume that the vector $\mathbf{x}$ of pseudo-symbols in (1) is Gaussian, the GSML estimate is given my the minimizer of [4]:

$$\mathbf{L}_{\text{GSML}}\left(\mathbf{r}/\lambda\right) = \text{tr}\left(\mathbf{R}_\lambda^{-1}\widehat{\mathbf{R}}\right) \tag{7}$$

where $\mathbf{R}_\lambda = \mathbf{A}_\lambda\mathbf{A}_\lambda^H + \sigma^2\mathbf{I}$ is (under the assumption of (5)) the expected covariance matrix of the received vector $\mathbf{r}$ that depends on the parameter λ. It is noted that the cost function (7) requires knowledge of the noise power, σ^2. In principle, the GSML function can be concentrated with respect to σ^2, by replacing σ^2 by its ML estimate: $\widehat{\sigma}^2 = \mathbf{r}^H\mathbf{P}_{\mathbf{A}_\lambda}^{\perp}\mathbf{r}/d$, where

$$\mathbf{P}_{\mathbf{A}_\lambda}^{\perp} = \mathbf{I} - \mathbf{A}_\lambda\left(\mathbf{A}_\lambda^H\mathbf{A}_\lambda\right)\mathbf{A}_\lambda^H \tag{8}$$

is the projector onto the orthogonal signal subspace spanned of dimension d. However, in that case, the resulting cost function becomes non-linear with respect to $\widehat{\mathbf{R}}$ (see [4], Eq. 4.43).

Remark 1 The application of the GSML approach, usually adopted in the array signal processing field, is questionable in the digital communication context, as the signals in $\mathbf{x}$ are discrete random variables (very far from being Gaussian). Anyway, $\mathbf{L}_{\text{GSML}}\left(\mathbf{r}/\lambda\right)$ is a reasonable cost function as it measures how well $\widehat{\mathbf{R}}$ fits into the signal subspace, giving more importance to the highest eigenvalues of the signal covariance matrix. Moreover, it can be proven ([4], corollary 4.1) that the Weighted Subspace Fitting (WSF) technique, which performs the *best weighted least-squares fit* between the estimated subspace and the model subspace, results in estimators which are large sample realizations of the GSML method. This is also in favor of the Gaussian assumption, even in the digital communications field.

3.3. Deterministic ML (DML)

Under the deterministic assumption, vector $\mathbf{x}$ is considered a parameter of the model that has to be estimated from $\mathbf{r}$. The joint ML function of λ and $\mathbf{x}$ is (to within irrelevant factors)

$$\mathbf{\Lambda}\left(\mathbf{r}/\lambda,\mathbf{x}\right) = \left(\mathbf{r} - \mathbf{A}_\lambda\mathbf{x}\right)^H\left(\mathbf{r} - \mathbf{A}_\lambda\mathbf{x}\right) \tag{9}$$

To avoid the joint search on λ and $\mathbf{x}$, we first solve for the minimum with respect to $\mathbf{x}$:

$$\widehat{\mathbf{x}}_{LS}\left(\mathbf{r},\lambda\right) = \left(\mathbf{A}_\lambda^H\mathbf{A}_\lambda\right)^{-1}\mathbf{A}_\lambda^H\mathbf{r} \tag{10}$$

Note that in the previous minimization, we are not assuming any constraint on the structure of $\mathbf{x}$ given in (4) (in particular that the pseudo-symbols $c_{m,i}$ are discrete). For that reason, $\widehat{\mathbf{x}}_{LS}\left(\mathbf{r},\lambda\right)$ is, in fact, the least-squares estimate of $\mathbf{x}$. Substituting (10) into (9) yields that the DML estimate is the minimizer of:

$$\mathbf{L}_{\text{DML}}\left(\mathbf{r}/\lambda\right) = \text{tr}\left(\mathbf{P}_{\mathbf{A}_\lambda}^{\perp}\widehat{\mathbf{R}}\right) \tag{11}$$

Remark 2 It is noted in [4] that the DML estimator is not efficient. In particular, its performance degrades rapidly in the lower to medium range of SNR. This is especially true in the case of high eigenvalue spread of the covariance matrix. In the problem of parameter estimation of CPM signals, high eigenvalue spreads occur for highly spectral efficient modulation formats, as a great number of pseudo-pulses with very dissimilar energies are obtained via Laurent expansion. Note that $\mathbf{L}_{\text{DML}}\left(\mathbf{r}/\lambda\right)$ (like $\mathbf{L}_{\text{GSML}}\left(\mathbf{r}/\lambda\right)$) measures how well $\widehat{\mathbf{R}}$ fits into the signal subspace, so the degradation in performance is due to the fact that those subspace signal components associated to small eigenvalues are very noisy. It should also be noted that the DML method requires that matrix $\mathbf{A}_\lambda$ be rank-deficient. The implication of this fact is that a lower bound exists in choosing the observation interval M, and this lower bound increases as the number of significant pseudo-pulses increases.

4. ESTIMATION-DIRECTED (ED) APPROACHES

We present a new approach to the problem which tries to solve the limitations of the ML methods explained above. The method is inspired on the philosophy of the DML, but introducing different statistical criteria. Starting from (9), we now assume that a general estimator of $\mathbf{x}$, $\widehat{\mathbf{x}}\left(\mathbf{r},\lambda\right)$, is available, resulting in the following compressed cost function:

$$\mathbf{L}\left(\mathbf{r}/\lambda\right) = \left(\mathbf{r} - \mathbf{A}_\lambda\widehat{\mathbf{x}}\left(\mathbf{r},\lambda\right)\right)^H\left(\mathbf{r} - \mathbf{A}_\lambda\widehat{\mathbf{x}}\left(\mathbf{r},\lambda\right)\right) \tag{12}$$

Clearly, the complexity-performance trade-off of the estimator of λ based on minimizing $\mathbf{L}\left(\mathbf{r}/\lambda\right)$, depends on the complexity and performance of the selected estimator $\widehat{\mathbf{x}}\left(\mathbf{r},\lambda\right)$. This estimator can be designed by using different criteria, as presented in the sequel.

4.1. Optimum-Detection ED

We can chose $\widehat{\mathbf{x}}(\mathbf{r},\lambda)=\widehat{\mathbf{x}}_{OD}(\mathbf{r},\lambda)$ as the optimum detector (OD) of $\mathbf{x}$ associated to the specific modulation format. The criterion coincides with the classical Decision-Directed (DD) approach: $\mathbf{L}_{\text{OD}-\text{ED}}(\mathbf{r}/\lambda)=\mathbf{L}_{\text{DD}}(\mathbf{r}/\lambda)$.

Remark 3 The ability of closed-loop DD schemes to converge depends on whether reliable decisions can be obtained even in the presence of (moderate) errors in the parameter λ, at the beginning of the convergence, which is usually not the case. Additionally, unless differential modulation schemes are used, phase recovery is a prerequisite to data recovery. Finally, DD methods usually introduce a decoding delay (especially in non-linear formats) that causes instabilities to the closed loop.

4.2. Least-Squares ED

We can chose $\widehat{\mathbf{x}}(\mathbf{r},\lambda)=\widehat{\mathbf{x}}_{LS}(\mathbf{r},\lambda)$ as the least-squares estimate of $\mathbf{x}$, as given in (10). This leads to the classical Deterministic ML (DML) approach explained above: $\mathbf{L}_{\text{LS}-\text{ED}}(\mathbf{r}/\lambda)=\mathbf{L}_{\text{DML}}(\mathbf{r}/\lambda)$.

4.3. Linear-Minimum-Mean-Squared-Error ED

With the motivation of avoiding very noisy estimates of the pseudo-symbols associated to low-energy pseudo-pulses, we can chose $\widehat{\mathbf{x}}(\mathbf{r},\lambda)=\widehat{\mathbf{x}}_{LMMSE}(\mathbf{r},\lambda)$ as the minimum-mean-squared-error estimate of $\mathbf{x}$, which is the well-known Wiener filter:

$$\widehat{\mathbf{x}}_{LMMSE}(\mathbf{r},\lambda)=\mathbf{A}_\lambda^H\left(\mathbf{A}_\lambda\mathbf{A}_\lambda^H+\sigma^2\mathbf{I}\right)^{-1}\mathbf{r} \tag{13}$$

Substituting (13) in (12) yields:

$$\mathbf{L}(\mathbf{r}/\lambda)=\mathbf{r}^H\mathbf{B}_\lambda^H\mathbf{B}_\lambda\mathbf{r} \tag{14}$$

where $\mathbf{B}_\lambda=\mathbf{I}-\mathbf{A}_\lambda\mathbf{A}_\lambda^H\left(\mathbf{A}_\lambda\mathbf{A}_\lambda^H+\sigma^2\mathbf{I}\right)^{-1}$. Adding and substracting $\sigma^2\mathbf{I}$ to the first outer product $\mathbf{A}_\lambda\mathbf{A}_\lambda^H$ yields $\mathbf{B}_\lambda=\sigma^2\mathbf{R}_\lambda^{-1}$, and then from (14) we obtain:

$$\mathbf{L}_{\text{MMSED-ED}}(\mathbf{r}/\lambda)=\text{tr}\left(\left(\mathbf{R}_\lambda^H\mathbf{R}_\lambda\right)^{-1}\widehat{\mathbf{R}}\right) \tag{15}$$

Remark 4 The new cost function $\mathbf{L}_{\text{MMSED-ED}}(\mathbf{r}/\lambda)$ (like $\mathbf{L}_{\text{GSML}}(\mathbf{r}/\lambda)$ and $\mathbf{L}_{\text{DML}}(\mathbf{r}/\lambda)$) measures how well $\widehat{\mathbf{R}}$ fits into the signal subspace, giving more importance to the highest eigenvalues of the signal covariance matrix, in proportion to its squared magnitude. It is worth mentioning that this is the underlying philosophy of the Weighted Subspace Fitting (WSF) criterion [4]. However, in the new cost function this weighting is performed without a explicit eigendecomposition of the sample covariance matrix.

5. RELATIONSHIP AMONG GSML, DML AND LMMSE-ED

To understand the relation among the different approaches, we resort to the matrix inversion lema, $\mathbf{R}_\lambda^{-1}=\sigma^{-2}\mathbf{I}-\sigma^{-2}\mathbf{A}_\lambda\left(\mathbf{A}_\lambda^H\mathbf{A}_\lambda+\sigma^2\mathbf{I}\right)^{-1}\mathbf{A}_\lambda^H$, and examine the following limiting cases.

Case 1: very large or very small σ^2. We obtain the following limiting values:

$$\sigma^2\mathbf{R}_\lambda^{-1}\underset{\sigma^2\to 0}{\to}\mathbf{P}_{\mathbf{A}_\lambda}^\perp$$
$$\sigma^4\mathbf{R}_\lambda^{-1}\underset{\sigma^2\to\infty}{\to}-\left(\mathbf{A}_\lambda\mathbf{A}_\lambda^H\right)^*+\sigma^2\mathbf{I}$$
$$\sigma^4\left(\mathbf{R}_\lambda^H\mathbf{R}_\lambda\right)^{-1}\underset{\sigma^2\to 0}{\to}\mathbf{P}_{\mathbf{A}_\lambda}^\perp$$
$$\sigma^8\left(\mathbf{R}_\lambda^H\mathbf{R}_\lambda\right)^{-1}\underset{\sigma^2\to\infty}{\to}-2\sigma^2\left(\mathbf{A}_\lambda\mathbf{A}_\lambda^H\right)^*+\sigma^4\mathbf{I}$$

As the terms $\sigma^2\mathbf{I}$ and $\sigma^4\mathbf{I}$ do not depend on λ, the DML, GSML and LLMSE-ED cost functions become equivalent (to within irrelevant scale factors). Note that in the case of very large σ^2, the estimate becomes the maximizer of $\text{tr}\left(\mathbf{R}_\lambda^*\widehat{\mathbf{R}}\right)$, which is a inner product between the estimated and expected second order statistics.

Case 2: orthonormal pulse shaping, $\mathbf{A}_\lambda^H\mathbf{A}_\lambda=E_p\mathbf{I}$. It is not difficult to see that all three cost functions become equivalent at any SNR. This condition holds in the case of linear modulations ($J=1$), as then $g_0(t)$ is usually designated such that $g_0*g_0(kT)=0$ for $k\neq 0$ (ISI-free condition).

Case 3: the eigenvalues of $\mathbf{A}_\lambda\mathbf{A}_\lambda^H$, are all equal to γ^2. Then, the LMMSE-ED and GSML are equivalent. This can be seen by noting that the GSML solution for an assumed noise power σ_1^2 depends only on the ratio $\left(\gamma^2+\sigma_1^2\right)/\sigma_1^2$, while the LLMSE-ED solution for an assumed noise power σ_2^2 depends only on the ratio $\left[\left(\gamma^2+\sigma_2^2\right)/\sigma_2^2\right]^2$. Choosing $\sigma_1^2=\sigma_2^2/\left(\gamma^2/\sigma_2^2+2\right)$ leads to identical ratios and, then, identical solutions.

Remark 5 From the above comments we conclude that it is worth exploring the possible advantages of the LMMSE-ED approach over the GSML only in the case of high eigenvalue spreads, and in the medium range of SNR. This is precisely the situation in the case of binary CPM signal of highly spectral efficiency.

6. FREQUENCY ERROR DETECTOR

As a case study, we derive a practical Frequency Error Detector (FED)[2]. Let us consider the general cost function $\mathbf{L}(\mathbf{r}/f)=\text{tr}\left(\mathbf{W}_f\widehat{\mathbf{R}}\right)$, where the weighting matrix is $\mathbf{W}_f=\mathbf{R}_f^{-1}$ (for GSML) or $\mathbf{W}_f=\left(\mathbf{R}_f^H\mathbf{R}_f\right)^{-1}$ (for LLMSE-ED). A measure of the frequency error frequency error, ε_f, can be obtained by computing the gradient of the cost function:

$$\varepsilon_f=\mathbf{L}'(\mathbf{r}/f)=\text{tr}\left(\mathbf{W}_f'\widehat{\mathbf{R}}\right)$$

For the problem of frequency estimation, the dependence of $\mathbf{R}_f$ on f is very simple: $\mathbf{R}_f=\mathbf{R}_0\odot\mathbf{E}_f$, where $[\mathbf{E}_f]_{p,q}=e^{j2\pi f(p-q)}$. After some manipulation, it is not difficult to show that $\varepsilon_f=\text{Im tr}\left(\mathbf{W}_0\left(\mathbf{X}\odot\widehat{\mathbf{R}}\odot\mathbf{E}_f^*\right)\right)$, where $[\mathbf{X}]_{p,q}=2\pi(p-q)$ and $\odot$ denotes the Hadamard matrix product.

[2] One can use a similar procedure to derive a practical Timing Error Detector (TED). This is omited for space reasons.

Notice that $\widehat{\mathbf{R}} \odot \mathbf{E}_f^* = \widehat{\mathbf{R}}_s$ represents the sample covariance matrix of the frequency shifted signal, $s(t) = r(t)\, e^{-j2\pi f t}$. Then, expressing $\mathbf{X}$ as $\mathbf{X} = \mathbf{u}\mathbf{v}^T - \mathbf{v}\mathbf{u}^T$, where $\mathbf{u} = [1, 1, \ldots, 1]^T$ and $\mathbf{v} = [0, 1, \ldots, M-1]^T$, and taking into account the synchronized averaging expression (6), we obtain[3]:

$$\varepsilon_f = 2 \sum_{k=0}^{N_s - 1} \mathrm{Im}\left(\mathbf{s}_{kT}^H \mathbf{W}_o \left(\mathbf{v} \odot \mathbf{s}_{kT}\right)\right) \qquad (16)$$

where $\mathbf{W}_o = \mathbf{R}_0^{-1}$ for GSML and $\mathbf{W}_o = \left(\mathbf{R}_0^H \mathbf{R}_0\right)^{-1}$ for LMMSE-ED. It is seen in (16) that the frequency error is computed as the sum of N_s terms, each one representing instantaneous (symbol-by-symbol) frequency error estimates obtained from a FED.

7. RESULTS

Figure 1 depicts the performance of the GSML (dotted) and LMMSE-ED estimators (solid), for different scenarios: A) assumed E_s/N_o of 10dB (GSML) and 5.2dB (LMMSE-ED); B) assumed E_s/N_o of 20dB (GSML) and 13dB (LMMSE-ED); C) assumed E_s/N_o of 30dB (GSML) and 21dB (LMMSE-ED). The criterion for the selection of the assumed E_s/N_o (which can be different for each method), has been that both methods yield (approximately) the same performance at low SNR. The following observations can be made:

- For a given assumed E_s/N_o for each method, the performance of the LMMSE-ED method is slightly superior at any (actual) E_s/N_o.
- When the assumed E_s/N_o is low (see scenario A), both methods yield the same performance (see section 5, Case 1).
- The assumed E_s/N_o determines the trade-off between the impact of additive noise and the impact of the self noise (or, equivalently, the floor jitter level of the curves).
- GSML and LMMSE-ED for scenario C yield a performance much better than the classical delay-and-multiply method (see [1] and references therein). The delay-and-multiply method was designed for MSK, and its performance degrades significantly when the input signal is GMSK, as it can be seen in the figure. Moreover, it performs fourth-order operations on the input signal while GSML and LMMSE-ED yield simpler quadratic schemes.
- The performance of the proposed schemes is still far from that predicted by the Modified Cramer-Rao Bound (MCRB) (see [1]). This fact tells us that we can expect an improvement on the performance by improving the quality of the general pseudo-symbol extractor $\widehat{\mathbf{x}}(\mathbf{r}, \lambda)$, for example, by relaxing the linear constraint (see [5]).

8. CONCLUSIONS AND FURTHER RESEARCH

The proposed Estimation-Directed approach constitutes a general framework for the derivation of simple algorithms

[3] $\mathbf{s}_t$ is defined similarly as (3).

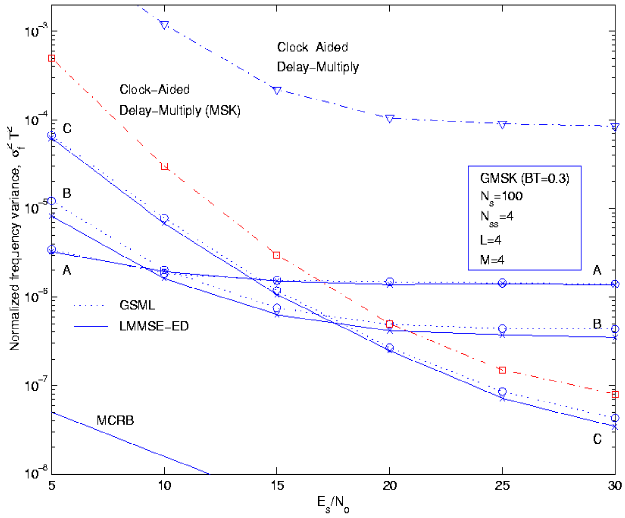

Figure 1: Performance of the GSML and LMMSE-ED estimators, for different assumed noise powers.

for timing and frequency estimation, and it can be applied to any modulation format admitting a Laurent decomposition (linear and binary CPM). The simplest algorithms are derived under the linear constraint of the general pseudo symbol extractor, and the obtained performance is near the same (indeed slightly better) as that obtained by assuming that the pseudo symbols are Gaussian. In fact, both assumption lead to quadratic schemes. Further research is being performed for relaxing the linear constraint to a linear plus third-order constrain, with the purpose to give a solid derivation of some ad-hoc fourth-order schemes derived in the literature for frequency and timing estimation.

9. REFERENCES

[1] Umberto Mengali and Aldo D'Andrea. *Synchronization Techniques for Digital Receivers*. Plenum Press, 1997.

[2] Gregori Vazquez and Jaume Riba. *Signal Processing Advances in Wireless and Mobile Communications*, by G.B. Giannakis, Y. Hua, P. Stoica and L. Tong, volume 2 (Trends in Single- and Multi-User Systems). Prentice Hall, July 2000. Chapter 9: Non-Data-Aided Digital Synchronization.

[3] Jaume Riba and Gregori Vázquez. "Non-Data-Aided Frequency Offset and Symbol Timing Estimation for Binary CPM: Performance Bounds". In *Proc. ICASSP'2000, Istambul (Turkey)*, pages pp. 2549–2552, June 2000.

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

[5] Xavier Villares, Gregori Vázquez, and Jaume Riba. "Fourth Order Non-Data-Aided Synchronization". In *Proc. ICASSP'2001, Salt Lake City (Utah)*, May 2001.