Conditional Maximum Likelihood Timing Recovery: Estimators and Bounds

Jaume Riba, Josep Sala, and Gregori Vázquez, *Senior Member, IEEE*

***Abstract*—This paper is concerned with the derivation of new estimators and performance bounds for the problem of timing estimation of (linearly) digitally modulated signals. The conditional maximum likelihood (CML) method is adopted, in contrast to the classical low-SNR unconditional ML (UML) formulation that is systematically applied in the literature for the derivation of non-data-aided (NDA) timing-error-detectors (TEDs). A new CML TED is derived and proved to be self-noise free, in contrast to the conventional low-SNR-UML TED. In addition, the paper provides a derivation of the conditional Cramér–Rao Bound (CRB_c), which is higher (less optimistic) than the modified CRB (MCRB) [which is only reached by decision-directed (DD) methods]. It is shown that the CRB_c is a lower bound on the asymptotic statistical accuracy of the set of consistent estimators that are quadratic with respect to the received signal. Although the obtained bound is not general, it applies to most NDA synchronizers proposed in the literature. A closed-form expression of the conditional CRB is obtained, and numerical results confirm that the CML TED attains the new bound for moderate to high E_s/N_o.**

***Index Terms*—Conditional maximum likelihood (CML), Cramér–Rao bound (CRB), pulse shaping, self-noise, synchronization, timing error detector (TED), timing recovery.**

I. INTRODUCTION

THE ESTIMATION of the timing epoch is a fundamental task of a digital receiver. Parameter estimation theory provides a means of obtaining practical estimators as well as performance lower limits on the variance of any unbiased estimator. The establishment of such limits is important because they provide benchmarks for evaluating the performance of actual estimators. In particular, Moeneclaey [1], [2] derived the Cramér–Rao bound (CRB) for the specific problem of timing estimation. More recently, it has been pointed out by D'Andrea *et al.* [3] that, in fact, the classical CRB for the timing, as well as the one derived for the carrier-frequency, belong to the class of so-called *modified* CRBs (MCRBs), which are generally lower than the true CRBs. The MCRB is much easier to compute and proves useful when, in addition to the parameter to be estimated, the observed data also depend on other unwanted parameters. The key problem is that in general, it is difficult to know in advance whether the MCRB is tight enough for use in practical applications [4].

Manuscript received March 16, 1999; revised December 15, 2000. This work was supported by Spanish Government Grants TIC98-0703, TIC99-0849, TIC2000-1025, and FIT-070000-2000-649 and by Generalitat de Catalunya Grant 2000 SGR-00083.

The authors are with the Department of Signal Theory and Communications, Universistat Politècnica de Catalunya, Barcelona, Spain (e-mail: jriba@gps.tsc.upc.es; alvarez@gps.tsc.upc.es; gregori@gps.tsc.upc.es).

Publisher Item Identifier S 1053-587X(01)02295-4.

In general, the observed signal from which the parameter of interest should be estimated also depends on a set of nuisance (unwanted) parameters. In the context of synchronization, nuisance parameters are the symbols, amplitude, and phase associated with the digital waveforms. Usually, symbols and phase are estimated in a second stage once the synchronization parameters have been acquired. In this way, the acquisition stage is simplified with respect to the more challenging goal of estimating these parameters jointly. Classical non data-aided (NDA) synchronization algorithms have been derived by treating the phase and data symbols as random variables [5]. When the above stochastic model is adopted, the ML function associated with the timing parameter is computed as the expectation of the joint ML, involving all parameters, with respect to the statistics assigned to the nuisance parameters. While the ML formulation in this way leads to practical algorithms (via approximations), insurmountable analytical obstacles appear in most cases when trying to compute the true CRB by means of the procedure above, due to the non-Gaussian nature of the signals. The MCRB provides a possible solution to this limitation, but the bound obtained departs from the true one in an unknown way.

In this paper, we resort to an entirely different approach for the computation of both the ML function and a performance bound, in which the symbols are modeled as *deterministic unknown* parameters. The approach followed is essentially the same as that widely applied in the context of sensor array processing (see [6] and references therein). When the deterministic approach is adopted, a *compressed* ML function is obtained by expressing the nuisance parameters as a function of the parameters of interest and the signal itself, instead of computing an *averaged* ML function. This formulation has usually been referred to as *conditional* (or *deterministic*) ML (CML) in contrast to the *unconditional* (or *stochastic*) ML (UML) adopted when signals are modeled as stochastic processes. The corresponding CRBs derived under these hypotheses are usually referred to as *conditional* CRB (CRB_c) and *unconditional* CRB (CRB_u), respectively.

The main contribution of this paper is the application of the CML principle to the timing estimation problem. We derive a new bound for the timing estimation problem that is more accurate than the MCRB and that it is not subject to the mathematical difficulties encountered in the computation of the true CRB. Moreover, it is shown that the new bound derived in this paper for timing estimation, which depends only on the second-order statistics of the symbol sequence, applies only to the set of timing estimators that are quadratic with respect to the received signal, which is a property shared by most NDA algorithms derived in the literature. Additionally, a practical (quadratic)

CML estimator for the timing, which has the property of being self-noise[1] free, is obtained. In this way, it is not necessary to resort to the classical *ad hoc* procedures [7], [8] proposed for reducing this noise. The CML principle has already been applied by the authors to frequency estimation of nonstaggered and staggered modulations in [9]. Its preliminary application to timing estimation can be found in [10]. Finally, a more recent work on frequency and timing estimation of continuous phase modulations (CPM) is presented in [11].

To our knowledge, the reason why the CML principle has never been applied to synchronization problems is that it requires a finite-dimensional representation of the received waveform. Classical synchronization algorithms have been derived employing continuous-time signal models, and for that reason, only the UML principle, along with the hypothesis of low-SNR, has been routinely adopted. While, in the context of sensor arrays, the finite representation appears naturally, the sampling operation is not essential in the formulation of synchronization problems, and the application of the CML principle is only possible if a discrete-time signal model is imposed from the beginning. This model, however, is adopted only for mathematical convenience, and the results obtained in this paper are general, irrespective of whether an analog or a digital receiver is used.

The paper is organized as follows. In Section II, a discrete-time signal model is proposed for the application of the ML principle. The proposed CML approach for timing recovery is explained in Section III, where a new TED and performance bound are derived. We pay special attention to the physical meaning and relationship between the low-SNR-UML and CML approaches in the specific context of timing estimation. Computer simulations comparing the performance of the CML and low-SNR-UML estimators are presented in Section IV, and conclusions are drawn in Section V.

II. Discrete-Time Signal Model for Timing Estimation and Classical ML Formulation

We assume that the received waveform has a complex envelope

$$r(t) = s_0(t) + w(t) \tag{1}$$

where $s_0(t)$ is the information-bearing signal, and $w(t)$ represents complex-valued white Gaussian noise with two-sided power spectral density $2N_o$. The signal $s_0(t)$ is modeled as follows:

$$s(t)_\tau = Ae^{j\theta}\sum_{i=0}^{L-1} c_i g(t-iT-\tau), \qquad s_o(t) = s(t)_{\tau=\tau_o} \tag{2}$$

where

- τ_o timing parameter to be estimated;
- τ timing estimate;
- $s(t)_\tau$ corresponding tentative model;
- θ signal phase;
- A signal amplitude;
- T symbol spacing;
- $\{c_i\}$ complex-valued symbols;
- L number of symbols in the observation interval;
- $g(t)$ (real-valued) signaling pulse of energy $E_g = \int_{-\infty}^{+\infty} g^2(t)\, dt$.

The set of unknown nuisance parameters includes the signal amplitude, the signal phase, and the data symbols and is denoted by the following vector

$$\mathbf{x}_o = A\mathbf{c}e^{j\theta_0} \tag{3}$$

where the data symbol vector $\mathbf{c}$ is given by

$$\mathbf{c} = [c_0 \cdots c_{L-1}].^T \tag{4}$$

For mathematical convenience in the formulation of the ML principle, we have considered a finite set of L symbols in the signal model (2). Notice that this is not restrictive as L can be made as large as desired. In the discrete model, the signal $r(t)$ is passed through an ideal antialiasing filter of bandwidth $f_s/2$ and sampled at a rate $f_s = 1/T_s$, where $T_s = T/N_{ss}$, where N_{ss} is an integer such that $N_{ss} > 1$ (to guarantee that f_s be above the Nyquist rate[2]). Then, (1) and (2) can be written in matrix notation as

$$\mathbf{r} = [r(0) \cdots r((M-1)T_s)]^T = \mathbf{A}_{\tau_0}\mathbf{x}_o + \mathbf{w} \tag{5}$$

where M is the number of nonzero samples of $r(t)$, which depends on the effective length of the signaling pulse, and

$$\begin{aligned}\mathbf{A}_\tau &= [\mathbf{a}_0(\tau) \cdots \mathbf{a}_{L-1}(\tau)]\\ \mathbf{a}_i(\tau) &= [g(-iT-\tau),\, g(T_s - iT - \tau) \cdots \\ &\qquad g((M-1)T_s - iT - \tau)]^T\\ \mathbf{w} &= [w(0) \cdots w((M-1)T_s)]^T.\end{aligned}$$

It is noted that the signal model (5) is a special case of that used in array signal processing in which each column of the transfer matrix $\mathbf{A}$ depends on a different parameter (typically, the direction-of-arrival of each signal). In our problem, the entire matrix $\mathbf{A}$ depends on a single parameter τ.

Although, in the case of infinite duration pulses, matrix $\mathbf{A}_\tau$ becomes semi-infinite, this does not modify the validity of the results presented through the paper. The matrix covariance of the noise vector after sampling is given by

$$\mathbf{C_w} = \sigma^2\mathbf{I} \tag{6}$$

where $\sigma^2 = 2N_o f_s$. From (5), we can formulate the joint ML function of τ and $\mathbf{x}$ as

$$\Lambda(\mathbf{r}|\tau, \mathbf{x}) = C\exp\left[-\frac{1}{\sigma^2}q(\mathbf{r}|\tau, \mathbf{x})\right] \tag{7}$$

where $q(\mathbf{r}|\tau, \mathbf{x})$ is the quadratic cost function defined as

$$q(\mathbf{r}|\tau, \mathbf{x}) = \|\mathbf{r} - \mathbf{A}_\tau\mathbf{x}\|^2 \tag{8}$$

and C is a positive constant that is irrelevant for maximization. In general, trying to estimate τ and $\mathbf{x}$ jointly from (7) is computationally intensive. Therefore, the goal is to obtain some cost function that is dependent only on τ and to derive NDA estimators from it, along with a bound on the performance

[1] This noise is referred to as pattern noise by some authors.

[2] $N_{ss} = 2$ suffices in the case of signals having an excess bandwidth less or equal than 100%.

of any estimator of τ independent of $\mathbf{x}$. To this end, two different approaches can be considered, which have been extensively studied in the context of array processing. They have been called the unconditional ML (UML) and the conditional ML (CML). The UML approach consists of modeling the nuisance parameters as random. Then, the UML function associated with the parameters of interest is computed as the expectation of the joint ML function with respect to the statistics of the nuisance parameters

$$\Lambda_u(\mathbf{r}|\tau) = E_{\mathbf{x}}\Lambda(\mathbf{r}|\tau, \mathbf{x}). \tag{9}$$

The UML estimator is given by the maximizer of $\Lambda_u(\mathbf{r}|\tau)$ with respect to τ. In general, the expectation $E_{\mathbf{x}}$ in (9) either poses insuperable obstacles or leads to a complicate cost function. Besides, maximization of $\Lambda_u(\mathbf{r}|\tau)$ requires knowledge of the operating SNR. Approximations of $\Lambda_u(\mathbf{r}|\tau)$ have been obtained (see [5]) under the hypothesis that the SNR is very low or very high. Moreover, the general CRB of the model

$$\mathrm{CRB}_u(\tau) = \frac{1}{E_{\mathbf{r}}\left\{\left|\dfrac{\partial \ln \Lambda_u(\mathbf{r}|\tau)}{\partial \tau}\right|^2\right\}} \tag{10}$$

which is a lower bound on the variance of any unbiased estimate, is even more difficult to compute in the general case. Only in some special cases, as for a Gaussian distribution of the nuisance parameters, does the derivation become straightforward. The MCRB [3]

$$\mathrm{MCRB}(\tau) = \frac{1}{E_{\mathbf{r},\,\mathbf{x}}\left\{\left|\dfrac{\partial \ln \Lambda(\mathbf{r}|\tau,\,\mathbf{x})}{\partial \tau}\right|^2\right\}} \le \mathrm{CRB}_u(\tau) \tag{11}$$

is also a lower bound on the variance of any unbiased estimate, and it reduces to

$$\mathrm{MCRB}(\tau) = \frac{1}{8\pi^2 L\xi}\,\frac{T^2}{E_s/N_o} \tag{12}$$

for the timing estimation problem, with $\xi = (\int_{-\infty}^{\infty} T^2 f^2 |G(f)|^2\, df / \int_{-\infty}^{\infty} |G(f)|^2\, df)$, $G(f)$ the Fourier transform of $g(t)$ and $E_s = (1/2)A^2 E|c_k|^2 E_g$ the symbol energy of the passband signal. In general, it cannot be known in advance whether the MCRB is tight enough for use in practical applications. This limitation motivates the new CML approach proposed in the next section.

III. Conditional ML Formulation

A. CML Function

In the conditional approach, the nuisance parameters are modeled as *deterministic*. Therefore, the vector $\mathbf{x}$ is substituted by that solution $\hat{\mathbf{x}}$ that unconstrainedly maximizes (7) for τ fixed. That is, no constraints whatsoever related to the distribution law of $\mathbf{x}$ are imposed[3] on the structure of $\hat{\mathbf{x}}$. Hence

$$\hat{\mathbf{x}} = \mathbf{A}_\tau^{\#}\mathbf{r} \tag{13}$$

[3]Contrary to the rule in the maximization of the joint ML function.

where $\mathbf{A}_\tau^{\#} = (\mathbf{A}_\tau^H\mathbf{A}_\tau)^{-1}\mathbf{A}_\tau^H$ denotes the pseudoinverse of matrix $\mathbf{A}_\tau$. Substituting $\mathbf{x}$ by $\hat{\mathbf{x}}$ in (8), we get the so-called compressed ML function or CML function

$$L_c(\mathbf{r}|\tau) = \ln\,\Lambda_c(\mathbf{r}|\tau) = \ln\,\Lambda(\mathbf{r}|\tau,\,\hat{\mathbf{x}}) = C_c\mathbf{r}^H\mathbf{A}_\tau\mathbf{A}_\tau^{\#}\mathbf{r} \tag{14}$$

where C_c is a positive constant irrelevant for maximization. Hence, the timing estimate resulting from the maximization of $L_c(\mathbf{r}|\tau)$ coincides with the τ-component of the joint estimate of $(\tau_0,\,\mathbf{x})$ obtained from the unconstrained maximization of $\Lambda(\mathbf{r}|\tau,\,\mathbf{x})$ in (7). It is seen that in the specific context of timing estimation, the CML function (14) is the cross-energy between the sampled matched filter $\mathbf{A}_\tau$ and the pseudoinverse filter $\mathbf{A}_\tau^{\#}$. In general, the filter $\mathbf{A}_\tau^{\#}$ is a zero-forcing equalizer because each component of $\mathbf{A}_\tau^{\#}\mathbf{r}$ is ISI-free for $\tau = \tau_o$: $\mathbf{A}_{\tau_o}^{\#}\mathbf{r} = \mathbf{x}_o + \mathbf{A}_{\tau_o}^{\#}\mathbf{w}$. In the particular case that the pulse $g(t)$ does not generate ISI, it holds that

$$\begin{aligned}\mathbf{A}_\tau^H\mathbf{A}_\tau &= \frac{E_g}{T_s}\,\mathbf{I}\\ \mathbf{A}_\tau^{\#} &= (\mathbf{A}_\tau^H\mathbf{A}_\tau)^{-1}\mathbf{A}_\tau^H = \frac{T_s}{E_g}\,\mathbf{A}_\tau^H\end{aligned} \tag{15}$$

and the CML function becomes simply the energy at the sampled matched filter output, irrespective of the symbol correlation. This is the result obtained with the low-SNR-UML formulation in the case of uncorrelated symbols. It is important to notice in this case that even when the two different approaches agree, a new performance bound and a new symbol-by-symbol timing error detector can be derived following the conditional approach (as proved further on).

B. Conditional CRB

We will now derive an expression for the conditional CRB by taking benefit of the research performed on the field of sensor array processing. The possibility to capitalize on the research in this field has been our main motivation in formulating a discrete-time signal model in Section II. Employing the conditional model assumption, Stoica and Nehorai [6] derived the *conditional* CRB (CRB_c). This bound, which in general cannot be reached, is applicable when no assumptions on the signal waveforms (i.e., the modulation symbols in the timing estimation problem) are made [12]. For the single parameter–single realization case, it can be expressed as[4]

$$\mathrm{CRB}_c(\tau) = \frac{\sigma^2}{2\mathbf{x}^H\mathbf{D}_\tau^H\mathbf{P}_{A_\tau}^{\perp}\mathbf{D}_\tau\mathbf{x}}. \tag{16}$$

Matrix $\mathbf{P}_A^{\perp}$, which is defined as

$$\mathbf{P}_{A_\tau}^{\perp} = \mathbf{I} - \mathbf{P}_{A_\tau},\ \mathbf{P}_{A_\tau} = \mathbf{A}_\tau\mathbf{A}_\tau^{\#} \tag{17}$$

is the projector onto the null space of $\mathbf{A}_\tau$, which is the space orthogonal to the signal subspace spanned by the columns of $\mathbf{A}_\tau$.

[4]The proof of (16) can be found in [13, App. E], where it is derived for the general model (5), considering that each column of the transfer matrix $\mathbf{A}$ depends on a different parameter (DOA). Expression (16) for the single parameter case can be obtained easily via application of the chain derivation rule to [13, Eq. (E.2d)], which has only impact on (E.8f) and (E.8g).

The derivative matrix $\mathbf{D}_\tau$ is defined as $d\mathbf{A}_\tau/d\tau$. For comparison purposes, with the classical expression (12), we write the previous bound in (16) as follows:[5]

$$\mathrm{CRB}_c(\tau) = \frac{1}{8\pi^2 L \xi'_L} \frac{T^2}{E_s/N_o} \tag{18}$$

where ξ'_L is an adimensional coefficient depending on the shape of $g(t)$ [similar to ξ in (12)], as well as on the symbol sequence properties

$$\xi'_L = \frac{T^2}{4\pi^2 E_g \sigma_x^2} T_s \frac{1}{L} \mathbf{x}^H \mathbf{D}_\tau^H \mathbf{P}_{A_\tau}^\perp \mathbf{D}_\tau \mathbf{x} \tag{19}$$

with $\sigma_x^2 = A^2 E|c_k|^2$. Let us focus first on the special case of a single symbol $L = 1$ of power $\sigma_x^2 = A^2|c_0|^2$. Then, we obtain

$$\begin{aligned}\xi'_1 = \xi &= \frac{T^2}{4\pi^2 E_g} T_s \left(|\mathbf{d}_0(\tau)|^2 - \frac{T_s |\mathbf{d}_0^H(\tau)\mathbf{a}_0(\tau)|^2}{E_g} \right) \\ &= \frac{T^2}{4\pi^2 E_g} T_s |\mathbf{d}_0(\tau)|^2 \\ &= \frac{\int_{-\infty}^{\infty} T^2 f^2 |G(f)|^2 \, df}{\int_{-\infty}^{\infty} |G(f)|^2 \, df}\end{aligned} \tag{20}$$

where it has been used that for real-valued pulses, the pulse and its derivative are orthogonal ($\mathbf{d}_0^H(\tau)\mathbf{a}_0(\tau) = 0$), and the last equality holds in virtue of Parseval's theorem. This means that for $L = 1$, $\mathrm{CRB}_c(\tau)_{\text{single pulse}} = \mathrm{MCRB}(\tau)$, because ξ'_1 coincides with the classical coefficient ξ defined in (12). In the next subsection, we examine the asymptotic case when $L \to \infty$.

C. Spectral Formulation of the Asymptotic $\mathrm{CRB}_c(\tau)$

In most applications, one is interested in obtaining a mean performance measure, independent of the specific realization of $\mathbf{x}$, as is the case in continuous mode. Such an indicator may be defined as the expected value of the conditional CRB, with respect to all realizations of $\mathbf{x}$, that is, $E_{\mathbf{x}}\mathrm{CRB}_c(\tau)$. However, $E_{\mathbf{x}}\mathrm{CRB}_c(\tau)$ is difficult to compute. Instead, a more manageable bound[6] can be formulated in view of Jensen's inequality $(1/E_{\mathbf{x}}[\xi'_L] \le E_{\mathbf{x}}[1/\xi'_L])$

$$\overline{\mathrm{CRB}}_c(\tau) = \frac{1}{8\pi^2 L E_{\mathbf{x}}[\xi'_L]} \frac{T^2}{E_s/N_o} \le E_{\mathbf{x}}\mathrm{CRB}_c(\tau). \tag{21}$$

Thus, the $\overline{\mathrm{CRB}}_c(\tau)$ is a valid lower bound on the variance of any consistent estimator when no assumptions on the modulation symbols are made [12]. The term $E_{\mathbf{x}}[\xi'_L]$ in (21) can be written as

$$E_{\mathbf{x}}[\xi'_L] = \frac{T^2}{4\pi^2 E_g \sigma_x^2} T_s \frac{1}{L} \mathrm{tr}(\mathbf{D}_\tau^H \mathbf{P}_{A_\tau}^\perp \mathbf{D}_\tau \boldsymbol{\Gamma}) \tag{22}$$

where

$$\boldsymbol{\Gamma} = E_{\mathbf{x}}[\mathbf{x}\mathbf{x}^H]. \tag{23}$$

On the other hand, a useful spectral characterization of $E_{\mathbf{x}}[\xi'_L]$ can be obtained by considering that L tends to infinity. In particular, it can be shown (see Appendixes A and B) that

$$\xi'_L \xrightarrow{(L\to\infty)} E_{\mathbf{x}}[\xi'_L] \xrightarrow{(L\to\infty)} \frac{T^2}{4\pi^2 E_g \sigma_x^2} \int_{-1/2}^{1/2} \Xi(e^{j2\pi F}) \, dF \tag{24}$$

with probability one, where

$$\Xi(e^{j2\pi F}) = \left(S_{dd}(e^{j2\pi F}) - \frac{|S_{da}(e^{j2\pi F})|^2}{S_{aa}(e^{j2\pi F})} \right) S_{xx}(e^{j2\pi F}). \tag{25}$$

The involved discrete spectra are defined as

$$\begin{aligned} S_{xx}(e^{j2\pi F}) &= \sum_{k=-\infty}^{+\infty} R_{xx}(k) e^{-j2\pi F k} \\ R_{xx}(k) &= E[x_i x_{i+k}^*] \end{aligned} \tag{26}$$

$$\begin{aligned} S_{aa}(e^{j2\pi F}) &= T_s \sum_{k=-\infty}^{+\infty} \overline{\mathrm{diag}}_k [\mathbf{A}_\tau^H \mathbf{A}_\tau] e^{-j2\pi F k} \\ &= \sum_{k=-\infty}^{+\infty} R_{gg}(kT) e^{-j2\pi F k} \\ &= \sum_{k=-\infty}^{\infty} \left| G\left(\frac{F+k}{T}\right)\right|^2 \end{aligned} \tag{27}$$

$$\begin{aligned} S_{dd}(e^{j2\pi F}) &= T_s \sum_{k=-\infty}^{+\infty} \overline{\mathrm{diag}}_k [\mathbf{D}_\tau^H \mathbf{D}_\tau] e^{-j2\pi F k} \\ &= \sum_{k=-\infty}^{+\infty} R_{g'g'}(kT) e^{-j2\pi F k} \\ &= \sum_{k=-\infty}^{\infty} \left(2\pi \frac{F+k}{T}\right)^2 \left| G\left(\frac{F+k}{T}\right)\right|^2 \end{aligned} \tag{28}$$

$$\begin{aligned} S_{da}(e^{j2\pi F}) &= T_s \sum_{k=-\infty}^{+\infty} \overline{\mathrm{diag}}_k [\mathbf{D}_\tau^H \mathbf{A}_\tau] e^{-j2\pi F k} \\ &= \sum_{k=-\infty}^{+\infty} R_{g'g}(kT) e^{-j2\pi F k} \\ &= \sum_{k=-\infty}^{\infty} j2\pi \frac{F+k}{T} \left| G\left(\frac{F+k}{T}\right)\right|^2 \\ &= -S_{ad}(e^{j2\pi F}) \end{aligned} \tag{29}$$

where $R_{gg}(t)$, $R_{g'g}(t)$, and $R_{g'g'}(t)$ are the temporal auto and cross-correlation of the pulses, and $\overline{\mathrm{diag}}_k$ represents the common element of the kth diagonal of the corresponding Toeplitz matrix.

[5] In a feed-back scheme, the parameter L is related to the equivalent noise bandwidth B_L as $L = 1/(2B_L T)$. Most authors write the CRB expressions as function of this bandwidth [5].

[6] It is noted that in [6], the bound defined in (21) is referred to as *asymptotic* CRB_c, meaning that the number of realizations of $\mathbf{x}$ tends to infinity. However, in this paper, by *asymptotic*, we understand that the dimension of $\mathbf{x}$ tends to infinity.

The consideration of large values of L leads to the definition of the asymptotic CRB_c as

$$\mathrm{CRB}_c^{as}(\tau) = \frac{1}{8\pi^2 L \xi'_\infty} \frac{T^2}{E_s/N_o}. \tag{30}$$

From (24), and with probability one, the following relation holds:

$$\frac{\mathrm{CRB}_c(\tau)}{\mathrm{CRB}_c^{as}(\tau)} \xrightarrow{(L\to\infty)} \frac{\overline{\mathrm{CRB}}_c(\tau)}{\mathrm{CRB}_c^{as}(\tau)} \xrightarrow{(L\to\infty)} 1 \tag{31}$$

which means that the $\mathrm{CRB}_c^{as}(\tau)$ becomes asymptotically (for large L) the true $\mathrm{CRB}_c(\tau)$ and that the $\mathrm{CRB}_c(\tau)$ tends to be independent of the specific realization of $\mathbf{x}$ with probability one. Thus, the lower bound in (21) becomes an asymptotic equality when $L \to \infty$.

The important point is that in contrast to the actual CRB_u (that is, the CRB_u computed by taking into account the true statistics of the signals), the CRB_c^{as} can be computed analytically. Moreover, Stoica and Nehorai showed [6] that, although in a general context the CRB_c cannot be attained, it converges to the Gaussian CRB_u (that is, the CRB_u computed under the hypothesis that all the signals are Gaussian) when the SNR increases or the dimension of the signal vector $\mathbf{r}$ increases. While in the context of sensor array processing the dimension of the signal vector is finite (i.e., equal to the number of sensors), in the context of timing estimation, the dimension of $\mathbf{r}$ increases in proportion to the number of symbols L. For that reason, and using (31), the new bound derived in this paper CRB_c^{as} converges asymptotically (for large data) to the same limit as the Gaussian CRB_u for large L. Moreover, the CML estimator attains this common bound under this hypothesis [6].

Of course, in presence of non-Gaussian signals (the case of data symbols in digital communications), the Gaussian CRB_u is no longer a lower bound on the variance of consistent estimators. In fact (see [14] corollary 2 and references therein), it applies only to the set of estimators that are asymptotically robust, i.e., those for which the actual distribution of the signal waveforms does not affect the asymptotic properties of the parameter estimates. It has been shown (see also [12, pp. 117, remark 4.2]) that the CML estimator is asymptotically robust. Moreover, the CML estimator (which is a quadratic estimator, as shown in the next subsection) attains asymptotically (for large data) the Gaussian CRB_u as well as the CRB_c^{as} (see [6]), as explained before. Therefore, the CML estimator is asymptotically the best estimator in the quadratic class. Hence, the new bound derived in this paper CRB_c^{as}, which depends only on the second-order statistics of the symbol sequence, is asymptotically (for large data) a valid lower bound on the variance of any consistent estimator that is quadratic with respect to the received signal; this is a property shared by most NDA algorithms derived in the literature.

In the particular case of square-root raised-cosine pulses with a given roll-off factor α, it is shown in Appendix C that (19) converges asymptotically with probability one to

$$\xi'_L \xrightarrow{(L\to\infty)} \begin{cases} \xi'_\infty = \dfrac{\alpha}{8}, & \alpha > 0 \\ \dfrac{1}{2\pi^2} \dfrac{\ln(L)+\gamma}{L}, & \alpha = 0 \end{cases} \tag{32}$$

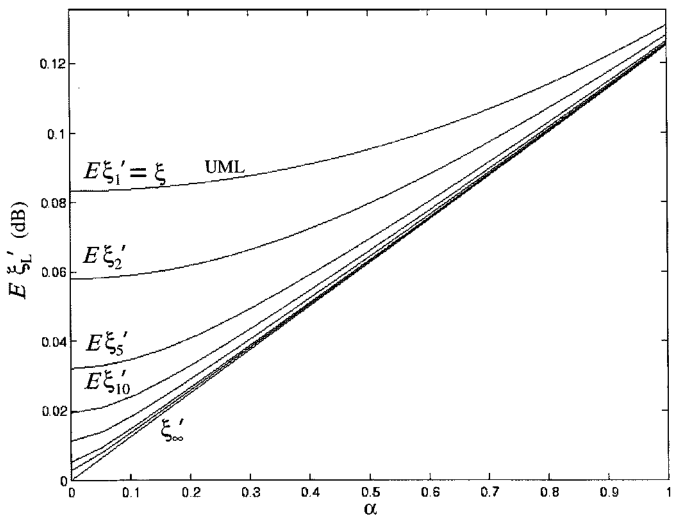


Fig. 1. $E\xi'_L$ as a function of the roll-off parameter α for $L = 1, 2, 5, 10, 20, 50, 100, \infty$.

where γ is Euler's constant. Substituting (19) into (18), we obtain

$$\mathrm{CRB}_c(\tau) \xrightarrow{(L\to\infty)} \begin{cases} \dfrac{1}{\pi^2 \alpha L} \dfrac{T^2}{E_s/N_o}, & \alpha > 0 \\ \dfrac{1}{8[\ln(L)+\gamma]} \dfrac{T^2}{E_s/N_o}, & \alpha = 0. \end{cases} \tag{33}$$

It its worth mentioning that the $\overline{\mathrm{CRB}}_c(\tau)$ shows a special behavior for $\alpha = 0$, decreasing as $\propto \ln^{-1}(L)$ for large L, instead of the classical dependence as $\propto L^{-1}$. Fig. 1 shows the evolution of $E\xi'_L$ for square-root raised-cosine pulses as a function of the roll-off parameter. It can be seen that all the curves tend to the value given by (32). Fig. 2 shows the evolution of $E\xi'_L$ as a function of L, which better illustrates the special behavior for $\alpha = 0$.

D. CML Timing Error Detector

In order to avoid the search in τ, the timing error detector can be formulated in terms of estimates of the (scalar) gradient of the CML function

$$\hat{\tau}_c = \overline{\tau} - \mu_c T_s \hat{\nabla}_c(\overline{\tau}) \tag{34}$$

where

- $\hat{\tau}_c$ timing estimate;
- $\overline{\tau}$ available previous estimate sufficiently close to the true parameter τ_o;
- μ_c corresponding step-size;
- $\hat{\nabla}_c(.)$ gradient of (14).

The scale factor T_s is introduced to cancel the dependence of the gradient magnitude on the sampling rate. The instantaneous gradient $\hat{\nabla}_c(\overline{\tau})$ is an estimate of $(1/L)\nabla_c(\overline{\tau})$, which is defined in (35). The factor $1/L$ normalizes the expected value of the block gradient $\nabla_c(\overline{\tau})$ to avoid $\mu_c \to 0$ when $L \to \infty$ to guarantee stability in updating the timing estimate. As shown further on, the way in which instantaneous estimates of the block gradient are produced allows the timing update procedure to operate at

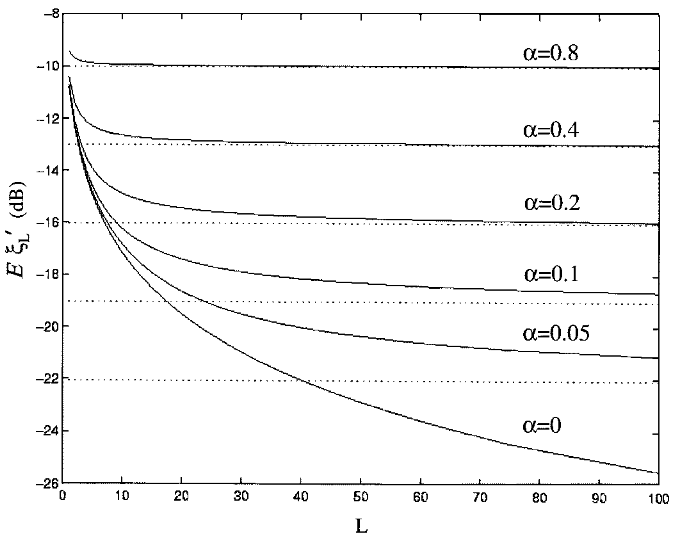


Fig. 2. $E\xi'_L$ as a function of L for different roll-off parameter, $\alpha = 0, 0.05, 0.1, 0.2, 0.4, 0.8$, along with their asymptotic values, ξ'_∞.

the symbol rate and not at the block rate. The conditional gradient $\nabla_c(\overline{\tau})$ has been obtained by Viberg *et al.* [15] within the more general context of sensor array processing. In the problem of timing estimation, all the columns of $\mathbf{A}_\tau$ are parameterized by the same τ, and the gradient can be expressed as

$$\nabla_c(\tau) = 2\text{Re}[\text{tr}(\mathbf{A}_\tau^{\#}\mathbf{r}\mathbf{r}^H\mathbf{P}_{A_\tau}^{\perp}\mathbf{D}_\tau)] = 2\text{Re}[\mathbf{r}^H\mathbf{P}_{A_\tau}^{\perp}\mathbf{D}_\tau\mathbf{A}_\tau^{\#}\mathbf{r}]. \tag{35}$$

It is seen that the gradient is estimated by measuring the cross-energy between the sampled signals at the output of matrices $\mathbf{A}_\tau^{\#}$ and $\mathbf{D}_\tau^H\mathbf{P}_{A_\tau}^{\perp}$. In general, consecutive rows of matrix $\mathbf{A}_\tau^{\#}$ do not differ from a simple time shift equal to N_{ss} components. (Contrarily, we find this simple structure in $\mathbf{A}_\tau$ and $\mathbf{D}_\tau$, having columns that are simply shifted versions of a single waveform). The same happens with the columns of matrix $\mathbf{P}_{A_\tau}^{\perp}\mathbf{D}_\tau$ because the projector operator breaks the original structure of matrix $\mathbf{D}_\tau$. In both cases, when matrix operators are formulated as filters, they become, in general, time-variant, except for matrix $\mathbf{A}_\tau^{\#}$ in the case of ISI-free pulse shaping, which becomes the matched filter. For that reason, we focus our attention on the derivation of asymptotic time-invariant filters from the gradient expression (35) with the purpose of obtaining a practical TED structure that is similar to that of the classical ML-oriented TED.

The asymptotic filters, which are defined as $g_c(t)$ and $d_c(t)$, are derived (see Appendix D) according to the following identification of the gradient in (35) with convolutions as $L \to \infty$:

$$\frac{T_s}{L}\nabla_c(\tau) = \frac{T_s}{L}\,2\text{Re}[\mathbf{r}^H\mathbf{P}_{A_\tau}^{\perp}\mathbf{D}\mathbf{A}_\tau^{\#}\mathbf{r}] \tag{36}$$

$$\xrightarrow{(L\to\infty)} \frac{1}{L}\sum_{k=0}^{L-1}\hat{\nabla}_c(\overline{\tau})_k \tag{37}$$

$$\hat{\nabla}_c(\overline{\tau})_k = 2\text{Re}[y'_a(k)y'^{*}_d(k)] \tag{38}$$

$$y'_a(k) = [r(t+\tau) * g_c(t)]_{t=kT} \tag{39}$$

$$y'_d(k) = [r(t+\tau) * d_c(t)]_{t=kT} \tag{40}$$

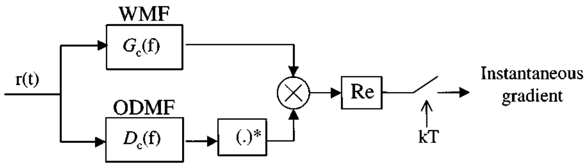


Fig. 3. Structure of the new CML TED.

where the following points can be noted.

i) In the limit, the normalized block gradient in (36) becomes, from (37), the expectation of the instantaneous gradient $\hat{\nabla}_c(\overline{\tau})_k$. To emphasize this fact, an asymptotic gradient $\nabla_c^{as}(\tau)$ can be defined such that with probability one

$$\nabla_c^{as}(\tau) = \lim_{L\to\infty}\left(\frac{T_s}{L}\nabla_c(\tau)\right) = E[\hat{\nabla}_c(\overline{\tau})_k]. \tag{41}$$

ii) From (34) and (38), the final feedback timing algorithm updates the timing estimate at the symbol rate as follows:

$$\hat{\tau}_c(k) = \hat{\tau}_c(k-1) - \mu_c\hat{\nabla}_c(\hat{\tau}_c(k-1))_k. \tag{42}$$

iii) The asymptotic filters $g_c(t)$ and $d_c(t)$ in (39) and (40) are obtained from the central columns of matrices $\mathbf{A}_\tau^{\#H}$ and $\mathbf{P}_{A_\tau}^{\perp}\mathbf{D}_\tau$ as $L \to \infty$, and the components of $\mathbf{A}_\tau^{\#}\mathbf{r}$ and of $\mathbf{D}_\tau^H\mathbf{P}_{A_\tau}^{\perp}\mathbf{r}$ asymptotically become the sampled output of $r(t+\tau) * g_c(t)$, $y'_a(k)$ and of $r(t+\tau) * d_c(t)$, $y'_d(k)$ at $t = kT$.

It is proven in Appendix D that these asymptotic filters can be expressed as

$$\text{WMF}, \quad G_c(f) = \frac{G^*(f)}{S_{aa}(e^{j2\pi fT})} \tag{43}$$

$$\text{ODMF}, \quad D_c(f) = G^*(f)\left(2\pi f - \frac{S_{ad}^*(e^{j2\pi fT})}{S_{aa}(e^{j2\pi fT})}\right). \tag{44}$$

These two filters will be referred to as whitened matched filter (WMF) and orthogonal derivative matched filter (ODMF). Observe that the ODMF equations (44) and (96), which are also valid for pulses subject to ISI, constitute a generalization of the one obtained by Moeneclaey in [8, eq. (97)]. Moeneclaey's solution was derived as a result of a constrained optimization problem in the ISI-free case using the *modified ML* principle, whereas in (44), the solution is general and derived directly from (conditional) ML arguments.

The asymptotic structure of the CML TED is shown in Fig. 3. The classical structure of the ML-oriented TED is the same, although the filters are the matched filter (MF) and the derivative matched filter (DMF). The main advantage of the new solution is that the ODMF does not generate self-noise because its output in the noiseless case is $\mathbf{D}_\tau^H\mathbf{P}_{A_\tau}^{\perp}\mathbf{A}_\tau\mathbf{x} = \mathbf{0}$ in the absence of timing error, as illustrated in Fig. 5 by the zero samples at the ODMF output. For that reason, the asymptotic ODMF filter is self-noise free, which is in contrast with the DMF.

It is important to notice that the asymptotic ODMF filter differs from the DMF filter even in the standard case of ISI-free pulse shaping $\mathbf{A}_\tau^{\#} = T_s\mathbf{A}_\tau^H/E_g$ when the CML and low-SNR-UML functions coincide within a scale factor to

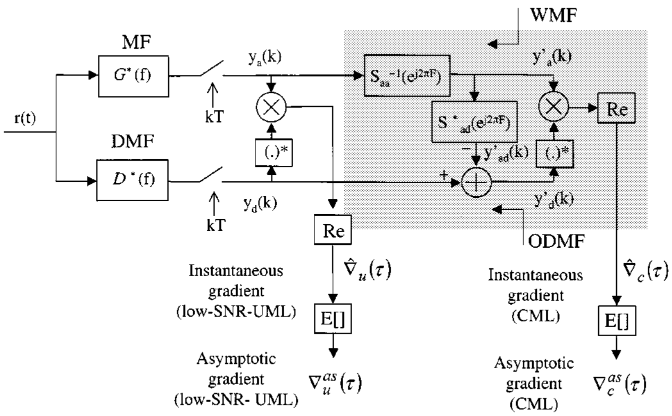


Fig. 4. Evaluation of instantaneous and asymptotic gradients for low-SNR-UML and CML as discrete filters on the decimated outputs of the MF and DMF. We have that $E[y'_a(k)y'^*_{ad}(k)] = 0$ in the ODMF scheme.

$\mathbf{r}^H\mathbf{A}_\tau\mathbf{A}_\tau^H\mathbf{r}$. This apparent contradiction is easily explained. The coincidence of the gradients $\nabla_u(\tau)$ (low-SNR-UML) $=$ $2\text{Re}[\mathbf{r}^H\mathbf{D}_\tau\mathbf{A}_\tau^H\mathbf{r}] \propto \nabla_c(\tau) = 2\text{Re}[\mathbf{r}^H\mathbf{P}_{A_\tau}^\perp\mathbf{D}_\tau\mathbf{A}_\tau^\#\mathbf{r}]$ does not imply the coincidence of the filters used to construct the asymptotic gradient (although the converse is true). Hence, although the asymptotic low-SNR-UML and CML filters differ, the asymptotic gradients $\nabla_u^{as}(\tau)$ and $\nabla_c^{as}(\tau)$ do not.

This is proven in more detail in Appendix D (self-noise cancellation in the ODMF). In summary, the substraction of the $S_{ad}(e^{j2\pi F})$ filter branch from the DMF output in Fig. 4 does not change the expected value of the instantaneous gradient in (38) on which the asymptotic gradient $\nabla_c^{as}(\tau)$ is based (41) so that a scheme operating on the cross-correlation between the outputs of the WMF and DMF would also be valid to implement an alternative[7] CML-TED. That is, the output of the $S_{ad}(e^{j2\pi F})$ filter branch $y'_{ad}(k)$ is not correlated with the WMF output. The motivation of a WMF/ODMF scheme over a WMF/DMF scheme stems from the self-noise cancellation property of the ODMF, as shown in (47) and Figs. 5 and 6. Then, the instantaneous gradient $\hat{\nabla}_c(\tau)$ is self-noise free for the optimum τ, and the TED need not rely on infinitely long time averages to cancel this noise contribution. This is specially advantageous in the medium to high range of E_s/N_o, where the dominant noise term in the lower (DMF) branch of the TED correlator is precisely self-noise. With respect to estimator performance, it should be noted that although CML and low-SNR-UML share the same asymptotic gradient, they yield different instantaneous gradients. As the timing algorithm makes use of the instantaneous rather than the asymptotic gradient, the tracking performance resulting from both methods will not be the same, as shown in Figs. 7 and 8.

For square-root raised-cosine pulses, it is not difficult to show (see Appendix C) that

$$
\begin{aligned}
\text{WMF}, &\quad G_c(f) = \frac{G^*(f)}{E_g} \\
\text{ODMF}, &\quad D_c(f) = G^*(f)W(fT)
\end{aligned}
\tag{45}
$$

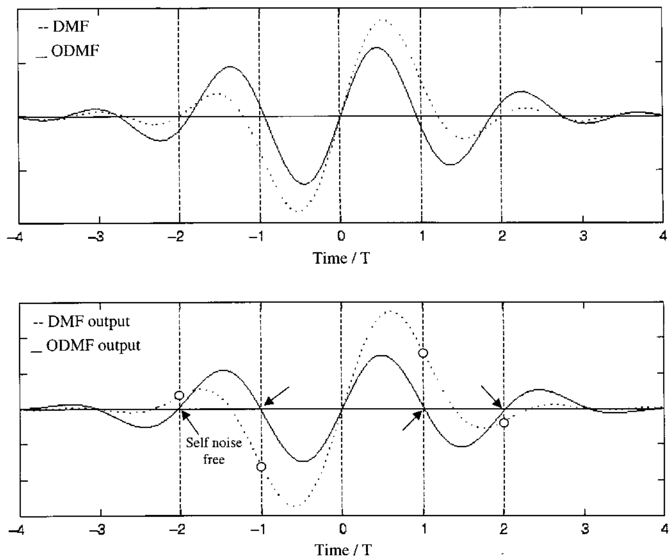


Fig. 5. Impulse response (up) and output to a single pulse (down) of the classical DMF and new ODMF for $\alpha = 1/2$. Self-noise is absent, as illustrated by the zero samples at the ODMF output.

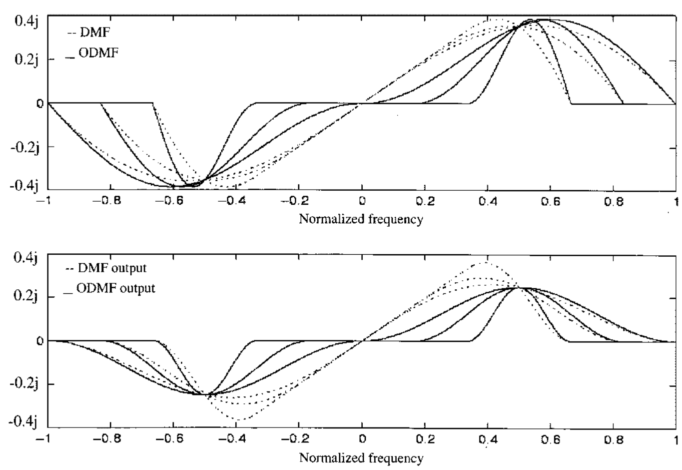


Fig. 6. Frequency response of the DMF and ODMF filters (up), as well as their outputs to a single pulse, for $\alpha = 1/3$, $\alpha = 2/3$, and $\alpha = 1$ (down). Self-noise at the ODMF output is absent, as illustrated by the mirror symmetry around half the symbol rate.

$$
W(F) = \begin{cases}
0, & |F| < \dfrac{1-\alpha}{2} \\[2ex]
\sin^2\left(\dfrac{\pi\left(F - \dfrac{1-\alpha}{2}\right)}{2\alpha}\right), & \dfrac{1-\alpha}{2} < F < \dfrac{1+\alpha}{2} \\[2ex]
-\sin^2\left(\dfrac{\pi\left(-F - \dfrac{1-\alpha}{2}\right)}{2\alpha}\right), & \dfrac{1-\alpha}{2} < -F < \dfrac{1+\alpha}{2}.
\end{cases}
$$

Fig. 6 shows the frequency response of the ODMF and of the ODMF output. Interestingly enough, the frequency response of the ODMF is different from zero only in the roll-off band, in contrast with the classical DMF response, which is zero only at

[7]We can generate an infinity of valid asymptotic filters by changing $S_{ad}(e^{j2\pi F})$ in Fig. 4 by any filter $H_{ad}(e^{j2\pi F})$, where $\text{Re}[H_{ad}(e^{j2\pi F})]$ is an odd function of F (Appendix C and D). Of course, the associated performance curves for finite L would vary.

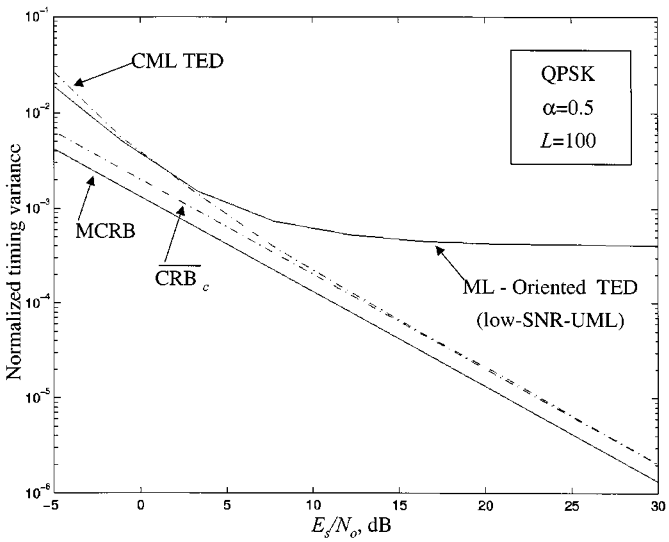


Fig. 7. Tracking performance of low-SNR-UML and CML TED's and Cramer–Rao bounds, for a rolloff factor of $\alpha = 0.5$.

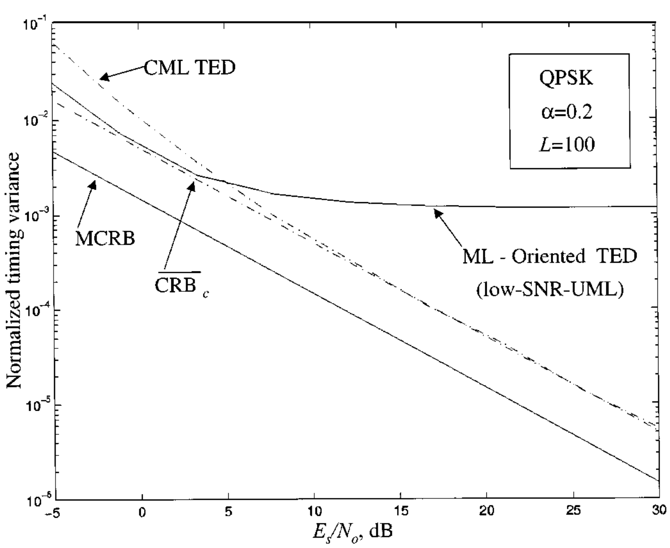


Fig. 8. Tracking performance of low-SNR-UML and CML TED's and CRBs for a roll-off factor of $\alpha = 0.2$.

the origin. Moreover, we can write the following closed-form expression of the ODMF output to a single pulse for $\overline{\tau} = \tau$

$$V_c(f) = G(f)D_c(f) = |G(f)|^2 W(fT) = \begin{cases} 0, & |f| < \frac{1-\alpha}{2T} \\ \frac{1}{4}\cos^2\left(\frac{\pi(fT-\frac{1}{2})}{\alpha}\right), & \frac{1-\alpha}{2T} < f < \frac{1+\alpha}{2T} \\ -\frac{1}{4}\cos^2\left(\frac{\pi(fT-\frac{1}{2})}{\alpha}\right), & \frac{1-\alpha}{2T} < -f < \frac{1+\alpha}{2T}. \end{cases} \quad (46)$$

The previous odd function, which is also depicted in Fig. 6, shows mirror symmetry around half the symbol rate, and therefore, its sampled output at the correct strobe instants is

$$v_c(kT) = \sum_{k=-\infty}^{\infty} V_c\left(f - \frac{k}{T}\right) = 0 \quad (47)$$

which constitutes an alternative illustration of the total self-noise cancellation. Finally, it is noted than the values of $\xi = \xi_1'$ and ξ_∞' are no more than the area under DMF and ODMF outputs to a single pulse respectively [see (78)].

Finally, the following conclusions can be drawn about the new CML TED.

- The CML gradient does not depend on the statistics of the data symbols.
- The symbol-by-symbol implementation of the CML TED based on the CML gradient involves two time-variant filters, the time-variant whitening matched filter ($\mathbf{A}_\tau^{\#}$), and the time-variant orthogonal derivative matched filter ($\mathbf{P}_{A_\tau}^{\perp}\mathbf{D}_\tau$), which differ from the classical matched (MF) and derivative (DMF) filters. $\mathbf{P}_{A_\tau}^{\perp}\mathbf{D}_\tau$ is time variant even in the standard case of ISI-free pulse shaping.
- The asymptotic time-invariant CML TED becomes self-noise free, and the two filters admit a closed-form expression in the frequency domain.

IV. Simulation Results

Numerical results are presented here to demonstrate the tracking performance of the CML TED compared with the low-SNR-UML TED (conventional ML-oriented TED). Figs. 7 and 8 show the normalized (with respect to T^2) timing variance as a function of E_s/N_o. Modulation is QPSK, and the overall channel response is Nyquist with rolloff 0.5 and 0.2, respectively. In both cases, a loop bandwidth of $5 \cdot 10^{-3}$ is chosen, which corresponds to an effective memory of $L = 100$ symbols. It is seen that the CML TED attains the asymptotic CRB_c at high E_s/N_o (in accordance with [6]), whereas the low-SNR-UML TED has a floor timing jitter due to self-noise. Self-noise is due to the overlapping in the time domain. This means that if the pulses were received one by one, the performance curve of the low-SNR-UML TED would have approached asymptotically to the MCRB curve. In contrast, the CML TED attains the asymptotic CRB_c, but it shows a variance penalty in the lower range of E_s/N_o. This penalty is higher for small excess bandwidth (rolloff), which is the case of a higher discrepancy of the MCRB from the asymptotic CRB_c.

There is another point to be remarked about the comparative performance results between the low-SNR-UML and the conditional ML timing estimators in the moderate-to-high range of E_s/N_o. The low-SNR-UML TED (without knowledge of the symbols) is derived under the hypothesis of low SNR, and is, in fact, an approximation to the *exact*, mathematically intractable, ML estimate (with SNR tending to zero, the exact and low-SNR-UML estimators would coincide). With increasing SNR, the low-SNR-UML estimator differs increasingly more from the exact ML estimator, and then, nothing can, in principle, be said about its performance. Observe now how this is reflected in the moderate-to-high range of E_s/N_o with the

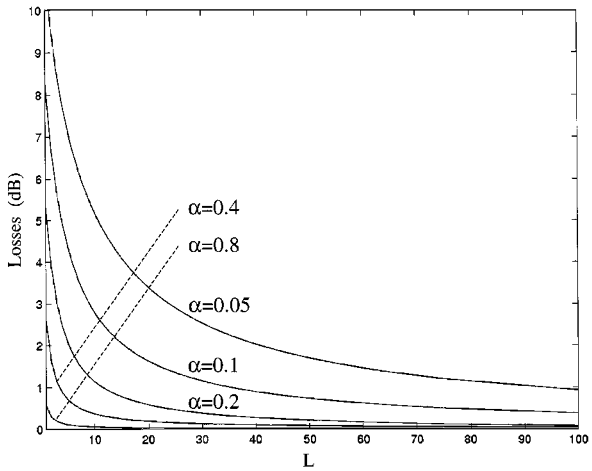


Fig. 9. Comparative performance of the approximate vs. exact CML estimator for uncorrelated QPSK. The losses are insensitive to the working E_s/N_o.

appearance of a performance floor. This has also been related before to the presence of self-noise in the low-SNR-UML TED (which is canceled in the CML scheme thanks to the zero-forcing operation $\hat{\mathbf{x}} = \mathbf{A}_\tau^{\#}\mathbf{r}$, although it is not the ML solution for the vector $\mathbf{x}_o$).

Finally, Fig. 9 compares the performance of the proposed (time-invariant) TED on the basis of the asymptotic gradient estimate in (38) with that of the exact (time-variant) CML estimator on the basis of the exact gradient in (35). This comparative analysis is useful to highlight the tradeoff between the complexity of the estimator and its estimation accuracy. It is seen that its loss is not significant for L sufficiently high (loop bandwidth small enough) as well as for large rolloff. The dependence of the performance loss on rolloff for fixed L is related to the fact that the central rows of the matrices involved in the exact CML take larger Ls to converge to their time-invariant structure. That is, pulses $g(t)$ with a small rolloff have longer tails.

V. Conclusions

A number of results developed in the field of sensor array processing have been translated to the problem of timing estimation of a (linearly) digitally modulated signal in the presence of unknown data symbols and phase. The CML principle has been considered on the basis of a discrete-time signal model. We have shown that the CML principle can be applied without the usual approximations and leads to new bounds and estimators.

Concerning the bounds, a new lower bound on the asymptotic variance of consistent quadratic timing estimators has been derived and formulated in the spectral domain. In combination with the classical MCRB, the new conditional bound gives us useful information about the performance of timing estimators: i) The MCRB is valid lower bound for all possible consistent estimators, although the NDA methods do not attain this bound in most cases; ii) the new conditional CRB is a more realistic bound that is asymptotically (for large data) attainable, although it is only valid for the subclass of quadratic NDA estimators. These estimators are often the most adequate for a practical implementation.

Concerning the estimators, a practical CML TED has been derived, which is self-noise free, in contrast with the low-SNR-UML TED that requires additional *ad hoc* techniques to cancel out this effect. We also find that the new CML TED attains the new lower bound for moderate-to-high E_s/N_o and sufficiently small loop bandwidth.

Appendix A
Consistency of the $\mathrm{CRB}_c(\tau)$

The purpose of this Appendix is the derivation of two propositions, which are to prove the following:

- the consistency of the $\mathrm{CRB}_c(\tau)$ using Proposition A.1 in Appendix B;
- an auxiliary result (Proposition A.2) necessary for Appendix D on asymptotic filters. This result is included in this Appendix because it can be readily derived using Proposition A.1.

The consistency of the $\mathrm{CRB}_c(\tau)$ is implied by

$$\lim_{L\to\infty} \xi'_L = \lim_{L\to\infty} E_{\mathbf{x}}[\xi'_L] \qquad \text{(with prob. one)} \tag{48}$$

[see (24)], which is proved using Proposition A.1 and the results in Appendix B. The spectral identity to express the corresponding asymptotic value

$$\lim_{L\to\infty} E_{\mathbf{x}}[\xi'_L] = \frac{T^2}{4\pi^2 E_g \sigma_x^2}\int_{-1/2}^{1/2} \Xi(e^{j2\pi F})\, dF \tag{49}$$

[see (24)] is proved in Appendixes B-D, based on Proposition B.3.

A. Proposition A.1

For some positive definite $L \times L$ matrix $\mathbf{M}^H\mathbf{M}$, if $\lim_{L\to\infty}(1/L)\mathrm{tr}(\mathbf{M}^H\mathbf{M}) < \infty$, the limit m_∞ defined from the second-order statistics[8] of $\mathbf{x}$ exists

$$m_\infty = \lim_{L\to\infty}\frac{1}{L}\,\mathrm{tr}(\mathbf{M}^H\mathbf{M}\boldsymbol{\Gamma}), \qquad \boldsymbol{\Gamma} = E(\mathbf{x}\mathbf{x}^H) \tag{50}$$

and if $\lambda_{\sup}(\mathbf{M}^H\mathbf{M}\boldsymbol{\Gamma})$ is a supremum bound on all eigenvalues of $\mathbf{M}^H\mathbf{M}\boldsymbol{\Gamma}$ when $L \to \infty$, then, with probability one, and for a single realization of process $\mathbf{x}$

$$\lim_{L\to\infty}\frac{1}{L}\,\mathbf{x}^H\mathbf{M}^H\mathbf{M}\mathbf{x} = m_\infty. \tag{51}$$

Proof: To this end, it suffices to show that $(1/L)\mathbf{x}^H\mathbf{M}^H\mathbf{M}\mathbf{x}$ is a random variable whose mean and co-variance tend to m_∞ (50) and zero, respectively. Therefore, the sequence of powers σ_L^2

$$\sigma_L^2 = E\left(\frac{1}{L}\,\mathbf{x}^H\mathbf{M}^H\mathbf{M}\mathbf{x} - \frac{1}{L}\,\mathrm{tr}(\mathbf{M}^H\mathbf{M}\boldsymbol{\Gamma})\right)^2 \tag{52}$$

should go to zero for $L\to\infty$. That is, $\sigma_\infty^2 = \lim_{L\to\infty}\sigma_L^2 \overset{(?)}{=} 0$. If this holds, and given that m_∞ is the limit of $(1/L)$ $\mathrm{tr}(\mathbf{M}^H\mathbf{M}\boldsymbol{\Gamma})$ (50), we must necessarily have that (51) is true. Using the Karthunen–Loéve expansion, we can express $\mathbf{x}$

[8]This proposition is also valid for any random vector $\mathbf{x}$ that is not necessarily the vector of modulation symbols. This allows us to use Propositions A.1 and A.2 in Appendix D.

as $\mathbf{x} = \mathbf{\Gamma}^{1/2}\mathbf{v}$, where $E\{\mathbf{v}\mathbf{v}^H\} = \mathbf{I}$. Defining matrix $\mathbf{\Omega}$ as $\mathbf{\Omega} = \mathbf{\Gamma}^{H/2}\mathbf{M}^H\mathbf{M}\mathbf{\Gamma}^{1/2}$ and expanding σ_L^2 in (52), yields

$$\begin{aligned}\sigma_L^2 &= E\left(\frac{1}{L}\,\mathbf{x}^H\mathbf{M}^H\mathbf{M}\mathbf{x}\right)^2 - \mathrm{tr}^2\left(\frac{1}{L}\,\mathbf{M}^H\mathbf{M}\mathbf{\Gamma}\right)\\ &= \frac{1}{L^2}(E(\mathbf{v}^H\mathbf{\Omega}\mathbf{v})^2 - \mathrm{tr}^2(\mathbf{\Omega})).\end{aligned}$$

Now, we may apply Proposition E.2 in Appendix E and

$$\sigma_\infty^2 = \lim_{L\to\infty}\frac{1}{L^2}(E(\mathbf{v}^H\mathbf{\Omega}\mathbf{v})^2 - \mathrm{tr}^2(\mathbf{\Omega})) \tag{53}$$

$$= (K-1)\lim_{L\to\infty}\frac{1}{L^2}\,\mathrm{tr}(\mathbf{\Omega}^2) \tag{54}$$

but $\mathrm{tr}(\mathbf{\Omega}^2) = \mathrm{tr}((\mathbf{M}^H\mathbf{M}\mathbf{\Gamma})^2) \le \lambda_{\sup}(\mathbf{M}^H\mathbf{M}\mathbf{\Gamma})\mathrm{tr}(\mathbf{M}^H\mathbf{M}\mathbf{\Gamma})$ and

$$\sigma_\infty^2 \le (K-1)\lambda_{\sup}(\mathbf{M}^H\mathbf{M}\mathbf{\Gamma})\lim_{L\to\infty}\frac{1}{L^2}\,\mathrm{tr}(\mathbf{\Omega}) = 0$$

for the finiteness of $\lim_{L\to\infty}(1/L)\mathrm{tr}(\mathbf{\Omega}) = m_\infty$ is a working assumption.

B. Consistency

Setting $\mathbf{M} = (T/2\pi\sigma_x)\sqrt{(T_s/E_g)}\mathbf{P}_{A_\tau}^{\perp}\mathbf{D}_\tau$ (Proposition A.1), with $\mathbf{P}_{A_\tau}^{\perp H}\mathbf{P}_{A_\tau}^{\perp} = \mathbf{P}_{A_\tau}^{\perp}$, we have from (19) that

$$\xi_L' = \frac{1}{L}\,\mathbf{x}^H\mathbf{M}^H\mathbf{M}\mathbf{x} \tag{55}$$

$$\mathbf{M}^H\mathbf{M} = \frac{T^2}{4\pi^2E_g\sigma_x^2}\,T_s(\mathbf{D}_\tau^H\mathbf{D}_\tau - \mathbf{D}_\tau^H\mathbf{A}_\tau\mathbf{A}_\tau^{\#}\mathbf{D}_\tau). \tag{56}$$

Consistency is implied by $m_\infty = \lim_{L\to\infty} E\xi_L' = \xi_\infty'$ (48) and (51). Therefore, from this particular definition of $\mathbf{M}$, we must prove that the conditions of Proposition A.1 are fulfilled. The first condition is that the limit m_∞ must exist. See proof in Appendix B-D. The second condition is that $\lim_{L\to\infty}(1/L^2)\mathrm{tr}(\mathbf{\Omega}^2)$ in (54) be zero [or that $\lambda_{\sup}(\mathbf{M}^H\mathbf{M}\mathbf{\Gamma})$ exists]. Using (54)

$$\begin{aligned}\sigma_\infty^2 &= (K-1)\lim_{L\to\infty}\frac{1}{L^2}\,\mathrm{tr}((\mathbf{M}^H\mathbf{M}\mathbf{\Gamma})^2)\\ &= (K-1)\frac{T^2}{4\pi^2E_g\sigma_x^2}\lim_{L\to\infty}\frac{1}{L}\\ &\quad\cdot\left(\lim_{L\to\infty}\frac{T_s}{L}\,\mathrm{tr}((\mathbf{D}_\tau^H\mathbf{D}_\tau - \mathbf{D}_\tau^H\mathbf{A}_\tau\mathbf{A}_\tau^{\#}\mathbf{D}_\tau)^2)\right)\end{aligned}$$

the decomposition into two limits is valid because

$$\lim_{L\to\infty}\frac{T_s^2}{L}\,\mathrm{tr}((\mathbf{D}_\tau^H\mathbf{D}_\tau - \mathbf{D}_\tau^H\mathbf{A}_\tau\mathbf{A}_\tau^{\#}\mathbf{D}_\tau)^2) = \int_{-1/2}^{1/2}\Xi^2(e^{j2\pi F})\,dF$$

is finite [trivial from (73) and Proposition B.3]. Hence, $\sigma_\infty^2 = 0$.

C. Proposition A.2

Generalization of Proposition A.1: With $\mathbf{R}_{21} = E(\mathbf{x}_2\mathbf{x}_1^H)$, we have

$$\lim_{L\to\infty}\frac{1}{L}\mathrm{Re}[\mathbf{x}_1^H\mathbf{M}^H\mathbf{M}\mathbf{x}_2] = \lim_{L\to\infty}\frac{1}{L}\mathrm{Re}[\mathrm{tr}(\mathbf{M}^H\mathbf{M}\mathbf{R}_{21})]. \tag{57}$$

Proof: Let us note that defining $\mathbf{x} = \mathbf{x}_1 - \mathbf{x}_2$

$$\begin{aligned}&\mathrm{Re}[\mathbf{x}_1^H\mathbf{M}^H\mathbf{M}\mathbf{x}_2]\\ &\quad= \tfrac{1}{2}\left(\mathbf{x}_1^H\mathbf{M}^H\mathbf{M}\mathbf{x}_1 + \mathbf{x}_2^H\mathbf{M}^H\mathbf{M}\mathbf{x}_2 - \mathbf{x}^H\mathbf{M}^H\mathbf{M}\mathbf{x}\right).\end{aligned}$$

All three terms may go by Proposition A.1 and the associated conditions, and (57) is proved

$$\begin{aligned}&\lim_{L\to\infty}\frac{1}{L}\,\mathrm{Re}[\mathbf{x}_1^H\mathbf{M}^H\mathbf{M}\mathbf{x}_2]\\ &\quad= \lim_{L\to\infty}\frac{1}{L}\left[\frac{1}{2}(\mathbf{x}_1^H\mathbf{M}^H\mathbf{M}\mathbf{x}_1 + \mathbf{x}_2^H\mathbf{M}^H\mathbf{M}\mathbf{x}_2 - \mathbf{x}^H\mathbf{M}^H\mathbf{M}\mathbf{x})\right]\\ &\quad= \lim_{L\to\infty}\frac{1}{L}\left[\mathrm{tr}\left(\mathbf{M}^H\mathbf{M}\,\frac{1}{2}\,E(\mathbf{x}_1\mathbf{x}_1^H + \mathbf{x}_2\mathbf{x}_2^H - \mathbf{x}\mathbf{x}^H)\right)\right]\\ &\quad= \lim_{L\to\infty}\frac{1}{L}\,\mathrm{Re}[\mathrm{tr}(\mathbf{M}^H\mathbf{M}\mathbf{R}_{21})].\end{aligned}$$

Appendix B
Spectral Characterization of ξ_∞'

This Appendix proves that $\lim_{L\to\infty} E_{\mathbf{x}}[\xi_L'] = (T^2/4\pi^2 E_g\sigma_x^2)\int_{-1/2}^{1/2}\Xi(e^{j2\pi F})\,dF$ [see (24)] in Section D of this Appendix, using Proposition B.3. Propositions B.1 and B.2 are only auxiliary to the proof of Proposition B.3, which will also be used in Appendix D.

A. Proposition B.1

$$\lim_{L\to\infty}\frac{1}{L}\,\mathrm{tr}\left(\prod_{k=1}^{N}\mathbf{R}_k\right) = \int_{-1/2}^{+1/2}\prod_{k=1}^{N}S_k(e^{j2\pi F})\,df \tag{58}$$

where $\mathbf{R}_i$ are N Toeplitz $L\times L$ matrices whose diagonals are

$$\mathrm{diag}_l[\mathbf{R}_k] = r_k(l)\mathbf{1}_{L-|l|} \qquad l = -(L-1),\,\ldots,\,(L-1) \tag{59}$$

and $S_k(e^{j2\pi F})$ are the discrete Fourier transforms of real sequences $r_k(l)$, which are given by

$$S_k(e^{j2\pi F}) = \lim_{L\to\infty}\sum_{l=-L}^{L} r_k(l)e^{-j2\pi Fl}. \tag{60}$$

Proof: We can express the trace as follows:

$$\begin{aligned}\mathrm{tr}\left(\prod_{k=1}^{N}\mathbf{R}_k\right) = \sum_{\substack{\{i_k\}=0\\ k=1\cdots N}}^{L-1} & r_1(i_N - i_1)r_2(i_1 - i_2)\cdots r_{N-1}\\ &\cdot(i_{N-2} - i_{N-1})r_N(i_{N-1} - i_N).\end{aligned} \tag{61}$$

Using the inverse discrete Fourier transform, we have

$$\lim_{L\to\infty}\frac{1}{L}\operatorname{tr}\left(\prod_{k=1}^{N}\mathbf{R}_k\right) = \lim_{L\to\infty}\frac{1}{L}\sum_{\substack{\{i_k\}=0\\k=1\cdots N}}^{L-1}\left(\int_{-\infty}^{\infty}S_1(e^{j2\pi F_1})e^{j2\pi F_1(i_N-i_1)}\,dF_1\right) \cdot\left(\int_{-\infty}^{\infty}S_2(e^{j2\pi F_2})e^{j2\pi F_2(i_1-i_2)}\,dF_2\right)\cdots \left(\int_{-\infty}^{\infty}S_{N-1}(e^{j2\pi F_{N-1}})e^{j2\pi F_{N-1}(i_{N-2}-i_{N-1})}\,dF_{N-1}\right) \cdot\left(\int_{-\infty}^{\infty}S_N(e^{j2\pi F_N})e^{j2\pi F_N(i_{N-1}-i_N)}\,dF_N\right). \quad (62)$$

In a more compact form, we can write

$$\lim_{L\to\infty}\frac{1}{L}\operatorname{tr}\left(\prod_{k=1}^{N}\mathbf{R}_k\right) = \lim_{L\to\infty}\frac{1}{L}\sum_{\substack{\{i_k\}=0\\k=1\cdots N}}^{L-1}\int_{\substack{\{F_{k'}\}=-\infty\\k'=1\cdots N}}^{\infty}S_1(e^{j2\pi F_1})S_2(e^{j2\pi F_2})\cdots S_{N-1}(e^{j2\pi F_{N-1}})S_N(e^{j2\pi F_N}) e^{j2\pi[F_1(i_N-i_1)+F_2(i_1-i_2)+\cdots+F_{N-1}(i_{N-2}-i_{N-1})+F_N(i_{N-1}-i_N)]} \cdot dF_1\,dF_2\cdots dF_{N-1}\,dF_N. \quad (63)$$

Under the assumption of absolute summability of $r_k(l)$ (uniform convergence of the Fourier transforms), the summation and integral operators in (63) can be interchanged

$$\lim_{L\to\infty}\frac{1}{L}\operatorname{tr}\left(\prod_{k=1}^{N}\mathbf{R}_k\right) = \int_{\substack{\{F_{k'}\}=-\infty\\k'=1\cdots N}}^{\infty}S_1(e^{j2\pi F_1})S_2(e^{j2\pi F_2})\cdots S_{N-1}(e^{j2\pi F_{N-1}}) \cdot S_N(e^{j2\pi F_N})\lim_{L\to\infty}\frac{1}{L}\sum_{\substack{\{i_k\}=0\\k=1\cdots N}}^{L-1} \cdot e^{j2\pi[F_1(i_N-i_1)+F_2(i_1-i_2)+\cdots+F_{N-1}(i_{N-2}-i_{N-1})+F_N(i_{N-1}-i_N)]} \cdot dF_1\,dF_2\cdots dF_{N-1}\,dF_N. \quad (64)$$

The limit inside the integral can be written as

$$\left(\lim_{L\to\infty}\frac{1}{L}\sum_{\{i_1\}=0}^{L-1}e^{j2\pi i_1(F_2-F_1)}\right) \cdot\left(\lim_{L\to\infty}\frac{1}{L}\sum_{\{i_2\}=0}^{L-1}e^{j2\pi i_2(F_3-F_2)}\right)\cdots \left(\lim_{L\to\infty}\frac{1}{L}\sum_{\{i_{N-1}\}=0}^{L-1}e^{j2\pi i_{N-1}(F_N-F_{N-1})}\right) \cdot\left(\lim_{L\to\infty}\frac{1}{L}\sum_{\{i_2\}=0}^{L-1}e^{j2\pi i_N(F_1-F_N)}\right). \quad (65)$$

Each of the previous N limits is equal to one iff their associated difference between frequencies is zero. Otherwise, the limits vanish. Therefore, the previous equation equals one iff $F_1 = F_2 = \cdots = F_{N-1} = F_N$. Then, (64) can be expressed with a single integral, as we wanted to prove

$$\lim_{L\to\infty}\frac{1}{L}\operatorname{tr}\left(\prod_{k=1}^{N}\mathbf{R}_k\right) = \int_{-\infty}^{\infty}S_1(e^{j2\pi F})S_2(e^{j2\pi F})\cdots S_{N-1}(e^{j2\pi F})S_N(e^{j2\pi F})\,dF. \quad (66)$$

B. Proposition B.2

$$\lim_{L\to\infty}\frac{1}{L}\operatorname{tr}\left(\mathbf{R}_1^{-1}\prod_{k=2}^{N}\mathbf{R}_k\right) = \int_{-1/2}^{+1/2}S_1^{-1}(e^{j2\pi F})\prod_{k=2}^{N} \cdot S_k(e^{j2\pi F})\,dF \quad (67)$$

where $\mathbf{R}_i$ are N Toeplitz $L\times L$ matrices as defined in Proposition B.1, and $\mathbf{R}_1$ must have $\operatorname{diag}_l[\mathbf{R}_1] = r_1(l)\mathbf{1}_{L-|l|} = r_1^*(-l)\mathbf{1}_{L-|l|}$.

Proof: To this end, we will first prove that any inverse of a Toeplitz matrix $\mathbf{R}$ can be expressed as

$$\mathbf{R}^{-1} = \varsigma^{-1}\sum_{m=0}^{+\infty}(\mathbf{I}-\varsigma^{-1}\mathbf{R})^m, \qquad \varsigma > \max_{1\le i\le L}[|\lambda_i(\mathbf{R})|] \quad (68)$$

where $\lambda_i(\mathbf{R})$ is the ith (nonzero) eigenvalue of the Toeplitz matrix $\mathbf{R}$. To prove this, let us construct $\overline{\mathbf{R}} = \varsigma^{-1}\mathbf{R}$. Then, if $\mathbf{R}$ displays Hermitian symmetry, its eigenvalue decomposition is given by $\mathbf{R} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^H$, where $\boldsymbol{\Lambda}$ is the diagonal eigenvalue matrix of $\mathbf{R}$, and $\mathbf{Q}^H = \mathbf{Q}^{-1}$. Therefore $\overline{\mathbf{R}} = \mathbf{Q}(\varsigma^{-1}\boldsymbol{\Lambda})\mathbf{Q}^H = \mathbf{Q}\overline{\boldsymbol{\Lambda}}\mathbf{Q}^H$ and all eigenvalues of $\overline{\mathbf{R}}$ fulfill $0<\lambda_i(\overline{\mathbf{R}})<1$. Then, for the unitary character of $\mathbf{Q}$ and the diagonal form of $\overline{\boldsymbol{\Lambda}}$, we can write

$$\begin{aligned}\overline{\mathbf{R}}^{-1} &= (\mathbf{I}-(\mathbf{I}-\overline{\mathbf{R}}))^{-1} = (\mathbf{I}-(\mathbf{I}-\mathbf{Q}\overline{\mathbf{\Lambda}}\mathbf{Q}^H))^{-1} \\ &= \mathbf{Q}(\mathbf{I}-(\mathbf{I}-\overline{\mathbf{\Lambda}}))^{-1}\mathbf{Q}^H \qquad (69)\\ &= \mathbf{Q}\sum_{m=0}^{+\infty}(\mathbf{I}-\overline{\mathbf{\Lambda}})^m\mathbf{Q}^H = \sum_{m=0}^{+\infty}(\mathbf{Q}(\mathbf{I}-\overline{\mathbf{\Lambda}})\mathbf{Q}^H)^m \\ &= \sum_{m=0}^{+\infty}(\mathbf{I}-\overline{\mathbf{R}})^m. \qquad (70)\end{aligned}$$

Substitution of $\overline{\mathbf{R}} = \varsigma^{-1}\mathbf{R}$ into (69) yields (68). Now, we may use Proposition B.1 because the trace expression tr $(\mathbf{R}_1^{-1}\prod_{k=2}^N \mathbf{R}_k)$ can be formulated in terms of powers of Toeplitz matrices as $\sum_{m=0}^{+\infty} \mathrm{tr}(\varsigma^{-1}(\mathbf{I}-\varsigma^{-1}\mathbf{R}_1)^m \prod_{k=2}^N \mathbf{R}_k)$. Thus, we have

$$\begin{aligned}&\lim_{L\to\infty}\frac{1}{L}\,\mathrm{tr}\left(\mathbf{R}_1^{-1}\prod_{k=2}^N \mathbf{R}_k\right)\\ &= \int_{-1/2}^{+1/2}\sum_{m=0}^{+\infty}\varsigma^{-1}(1-\varsigma^{-1}S_1(e^{j2\pi F}))^m\prod_{k=2}^N S_k(e^{j2\pi F})\,dF\\ &= \int_{-1/2}^{+1/2} S_1^{-1}(e^{j2\pi F})\prod_{k=2}^N S_k(e^{j2\pi F})\,dF.\end{aligned}$$

By $0 < \varsigma^{-1}S_1(e^{j2\pi F}) < 1$, we recover $S_1^{-1}(e^{j2\pi F})$, thus proving (67); the classical result $\lim_{L\to\infty}\max[\lambda_i(\mathbf{R}_1)] = \max[S_1(e^{j2\pi F})]$ assures that $\sum_{m=0}^{+\infty}\varsigma^{-1}(1-\varsigma^{-1}S_1(e^{j2\pi F}))^m$ converges.

C. Proposition B.3

$$\lim_{L\to\infty}\frac{1}{L}\,\mathrm{tr}\left(\prod_{k=1}^N \mathbf{R}_k^{l_k}\right) = \int_{-1/2}^{+1/2}\prod_{k=1}^N S_k^{l_k}(e^{j2\pi F})\,dF \qquad (71)$$

where $\mathbf{R}_k$ and $S_k^{l_k}(e^{j2\pi F})$ as in (59), (60), and $l_k \in \{-1, +1\}$.

Proof: (Straightforward) We proceed exactly as for Proposition B.2, decomposing each inverse into a power series and recomposing the associated spectrum within the integral via Proposition B.1 into $\prod_{k=1}^N S_k^{l_k}(e^{j2\pi F})$.

D. Proof of $\lim_{L\to\infty} E_{\mathbf{x}}[\xi'_L] = (T^2/4\pi^2 E_g\sigma_x^2)\int_{-1/2}^{1/2}\Xi(e^{j2\pi F})\,dF$

Using (17) and (19), we can write ξ'_L as

$$\xi'_L = \frac{T^2}{4\pi^2 E_g\sigma_x^2} \cdot T_s\left(\frac{1}{L}\,\mathrm{tr}(\mathbf{D}_\tau^H\mathbf{D}_\tau\mathbf{\Gamma}) - \frac{1}{L}\,\mathrm{tr}(\mathbf{D}_\tau^H\mathbf{A}_\tau\mathbf{A}_\tau^{\#}\mathbf{D}_\tau\mathbf{\Gamma})\right). \qquad (72)$$

The left term of (72) is obtained using Proposition B.3 for $N = 2$, where the Toeplitz matrices are $\mathbf{R}_1 = T_s\mathbf{D}_\tau^H\mathbf{D}_\tau$ and $\mathbf{R}_2 = \mathbf{\Gamma}$, and their spectrum $S_{dd}(e^{j2\pi F})$ and $S_{xx}(e^{j2\pi F})$ correspond to (28) and (26), respectively. The right term of (72) is obtained for $N = 4$, where the matrices are $\mathbf{R}_1 = T_s\mathbf{A}_\tau^H\mathbf{A}_\tau$, $\mathbf{R}_2 = T_s\mathbf{D}_\tau^H\mathbf{A}_\tau$, $\mathbf{R}_3 = T_s\mathbf{A}_\tau^H\mathbf{D}_\tau$, and $\mathbf{R}_4 = \mathbf{\Gamma}$, and their spectrum $S_{aa}(e^{j2\pi F})$, $S_{da}(e^{j2\pi F})$, $S_{ad}(e^{j2\pi F})$, and $S_{xx}(e^{j2\pi F})$ correspond to (27), (29), and (26), respectively. From $\Xi(e^{j2\pi F})$ in (25)

$$\lim_{L\to\infty}\frac{T_s}{L}\,\mathrm{tr}(\mathbf{D}_\tau^H\mathbf{D}_\tau\mathbf{\Gamma} - \mathbf{D}_\tau^H\mathbf{A}_\tau\mathbf{A}_\tau^{\#}\mathbf{D}_\tau\mathbf{\Gamma}) = \int_{-1/2}^{1/2}\Xi(e^{j2\pi F})\,dF \qquad (73)$$

proves the result. Neighboring columns of $\mathbf{A}_\tau$ and $\mathbf{D}_\tau$ are sampled versions of the pulses $g(t)$ and $g'(t)$ one symbol apart so that the components of each matrix $\mathbf{R}_i$ are the corresponding discrete cross-correlations evaluated at multiples of the symbol period T.

APPENDIX C
COMPUTATION OF THE ASYMPTOTIC ξ'_L FOR SQUARED-ROOT RAISED-COSINE PULSES

This Appendix proves (32). The frequency response of a squared-root raised-cosine pulse is

$$G(f) = \begin{cases} 1, & |f| \le \dfrac{1-\alpha}{2T} \\ \cos\left(\dfrac{\pi\left(|fT| - \dfrac{1-\alpha}{2}\right)}{2\alpha}\right), & \dfrac{1-\alpha}{2T} < |f| < \dfrac{1+\alpha}{2T} \\ 0, & |f| > \dfrac{1+\alpha}{2T} \end{cases} \qquad (74)$$

with $E_g = 1$. The spectra $S_{xx}(e^{j2\pi F})$ and $S_{aa}(e^{j2\pi F})$ involved in (26) and (27) can be expressed as

$$S_{xx}(e^{j2\pi F}) = \sigma_x^2, \qquad S_{aa}(e^{j2\pi F}) = \frac{1}{T_s}$$

where $S_{dd}(e^{j2\pi F})$ and $S_{da}(e^{j2\pi F})$ in (28) and (29) can be split into two frequency ranges

$$S_{dd}(e^{j2\pi F}) = \begin{cases} S_{dd}(e^{j2\pi F})_l, & |f| \le \dfrac{1-\alpha}{2} \\ S_{dd}(e^{j2\pi F})_h, & \dfrac{1-\alpha}{2} < |f| < \dfrac{1}{2} \end{cases}$$
$$S_{da}(e^{j2\pi F}) = \begin{cases} S_{da}(e^{j2\pi F})_l, & |f| \le \dfrac{1-\alpha}{2} \\ S_{da}(e^{j2\pi F})_h, & \dfrac{1-\alpha}{2} < |f| < \dfrac{1}{2} \end{cases} \qquad (75)$$

where

$$S_{dd}(e^{j2\pi F})_l = \frac{4\pi^2}{T_s T^2} F^2$$

$$\begin{aligned} S_{dd}(e^{j2\pi F})_h &= \frac{4\pi^2}{T_s T^2}\left[F^2 \cos^2\left(\frac{\pi\left(F - \frac{1-\alpha}{2}\right)}{2\alpha}\right)\right. \\ &\quad \left. + (F-1)^2 \cos^2\left(\frac{\pi\left(F - \frac{1+\alpha}{2}\right)}{2\alpha}\right)\right] \end{aligned}$$

$$|S_{da}(e^{j2\pi F})_l|^2 = \frac{4\pi^2}{T_s^2 T^3} F^2$$

$$\begin{aligned} |S_{da}(e^{j2\pi F})_h|^2 &= \frac{4\pi^2}{T_s^2 T^3}\left[F \cos^2\left(\frac{\pi\left(F - \frac{1-\alpha}{2}\right)}{2\alpha}\right)\right. \\ &\quad \left. + (F-1) \cos^2\left(\frac{\pi\left(F - \frac{1+\alpha}{2}\right)}{2\alpha}\right)\right]^2 \\ &= S_{dd}(e^{j2\pi F})_h + \frac{4\pi^2}{T_s^2 T^3} F(F-1) \\ &\quad \cdot \cos^2\left(\frac{\pi\left(F - \frac{1}{2}\right)}{\alpha}\right). \end{aligned} \tag{76}$$

After some algebraic manipulation, we obtain

$$\begin{aligned} \Xi(e^{j2\pi F}) &= S_{dd}(e^{j2\pi F}) - \frac{|S_{da}(e^{j2\pi F})|^2}{S_{aa}(e^{j2\pi F})} \\ &= \frac{1}{4}\cos^2\left(\pi\,\frac{F-1/2}{\alpha}\right) \end{aligned} \tag{77}$$

which is valid only in the range $(1-\alpha)/2 < F \le 1/2$. Finally

$$\xi'_\infty = 2\int_{(1-\alpha)/2}^{1/2} \frac{1}{4}\cos^2\left(\pi\,\frac{F-1/2}{\alpha}\right) dF = \frac{\alpha}{8}. \tag{78}$$

Now, the case of $\alpha = 0$ is considered. For $\mathbf{\Gamma} = \sigma_x^2 \mathbf{I}$ and $\mathbf{A}_\tau^\# = T_s \mathbf{A}_\tau^H$, the right term of (24) before taking the limit in $L \to \infty$ is

$$\sigma_x^2 \frac{1}{L}\,\mathrm{tr}([\mathbf{D}_\tau^H \mathbf{A}_\tau][\mathbf{A}_\tau^H \mathbf{D}_\tau]) = \sigma_x^2 \sum_{n=-(L-1)}^{L-1} \frac{L-|n|}{L} R_{g'g}^2(nT) \tag{79}$$

while the left term of (24) is

$$\sigma_x^2 \frac{1}{L}\,\mathrm{tr}(\mathbf{D}_\tau^H \mathbf{D}_\tau) = \sum_{n=-\infty}^{\infty} {g'}^2(nT). \tag{80}$$

For $\alpha = 0$, we have

$$g(t) = \sqrt{\frac{1}{T}}\,\frac{\sin(\pi t/T)}{\pi t/T}, \qquad R_{gg}(t) = \frac{1}{T} g(t)$$

$$R_{g'g}(t) = R'_{gg}(t) = \sqrt{\frac{1}{T}}\, g'(t)$$

$$g'(nT) = \begin{cases} 0 & \text{for } n = 0 \\ \sqrt{\frac{1}{T}}\,\frac{-(-1)^{|n|}}{n} & \text{for } n \neq 0. \end{cases} \tag{81}$$

Using the previous expressions, we can write, from (79)

$$\sigma_x^2 \sum_{n=-(L-1)}^{L-1} \frac{L-|n|}{L} R_{g'g}^2(nT) = 2\sigma_x^2 \sum_{n=1}^{L-1} \frac{1}{n^2} - \frac{2\sigma_x^2}{L} \sum_{n=1}^{L-1} \frac{1}{n}. \tag{82}$$

The left term of (82) converges faster than the right term and cancels out with (80). Then, using (19), and for large L, we can write

$$\xi'_L \xrightarrow{(L\to\infty)} \frac{1}{2\pi^2}\,\frac{1}{L}\sum_{n=1}^{L-1}\frac{1}{n} \xrightarrow{(L\to\infty)} \frac{1}{2\pi^2}\,\frac{\ln(L)+\gamma}{L} \tag{83}$$

where for $L \to \infty$, the harmonic series can be substituted by $\ln(L) + \gamma$, with $\gamma = 0.577\,21\ldots$, which is Euler's constant. That proves the result (32).

APPENDIX D
ASYMPTOTIC ODMF

In this Appendix, we prove the following.

- Subsection A: The identities in (38) and (41) as well as the expression for the WMF and the ODMF in terms of discrete filters (Fig. 4).
- Subsection B: The expression for the analog WMF and ODMF in (43) and (44) and in Fig. 3.
- Subsection C: Some properties related with the self-noise cancellation property of the WMF/ODMF scheme.

A. *Asymptotic Gradient and Filter Scheme*

Using the expressions for $\mathbf{P}_{A_\tau}^\perp$ and $\mathbf{A}_\tau^\#$ and Proposition A.2 in Appendix A, we can operate with the asymptotic gradient $\nabla_c^{as}(\tau)$ in (41) as follows:

$$\frac{1}{2}\nabla_c^{as}(\tau) = T_s \lim_{L\to\infty} \frac{1}{L}\mathrm{Re}[\mathbf{r}^H \mathbf{P}_{A_\tau}^\perp \mathbf{D}_\tau \mathbf{A}_\tau^\# \mathbf{r}] \tag{84}$$

$$\begin{aligned} &= T_s \lim_{L\to\infty} \frac{1}{L}\mathrm{Re}[\underline{\mathbf{r}^H \mathbf{D}_\tau}(\mathbf{A}_\tau^H \mathbf{A}_\tau)^{-1}\underline{\mathbf{A}_\tau^H \mathbf{r}}] \\ &\quad - T_s \lim_{L\to\infty} \frac{1}{L}\mathrm{Re}[\underline{\mathbf{r}^H \mathbf{A}_\tau}(\mathbf{A}_\tau^H \mathbf{A}_\tau)^{-1} \\ &\quad \cdot \underline{(\mathbf{A}_\tau^H \mathbf{D}_\tau)(\mathbf{A}_\tau^H \mathbf{A}_\tau)^{-1}\mathbf{A}_\tau^H \mathbf{r}}] \\ &= \lim_{L\to\infty} \frac{1}{L}\mathrm{Re}[\mathrm{tr}((T_s \mathbf{A}_\tau^H \mathbf{A}_\tau)^{-1} \\ &\quad \cdot \{E(T_s \mathbf{A}_\tau^H \mathbf{r})(T_s \mathbf{D}_\tau^H \mathbf{r})^H\})] - \lim_{L\to\infty}\frac{1}{L} \\ &\quad \cdot \mathrm{Re}[\mathrm{tr}((T_s \mathbf{A}_\tau^H \mathbf{A}_\tau)^{-1}(T_s \mathbf{A}_\tau^H \mathbf{D}_\tau)(T_s \mathbf{A}_\tau^H \mathbf{A}_\tau)^{-1} \\ &\quad \cdot \{E(T_s \mathbf{A}_\tau^H \mathbf{r})(T_s \mathbf{A}_\tau^H \mathbf{r})^H\})] \end{aligned} \tag{85}$$

with probability one, where we have underlined the terms identified with $\mathbf{x}_1^H$ and $\mathbf{x}_2$ in Proposition A.2, and $\mathbf{M}^H\mathbf{M} = (\mathbf{A}_\tau^H\mathbf{A}_\tau)^{-1}$. Now, $\mathbf{R}_{ad} = E(T_s\mathbf{A}_\tau^H\mathbf{r})(T_s\mathbf{D}_\tau^H\mathbf{r})^H$ is, by construction, the cross-correlation (Toeplitz) matrix between the MF output vector $\mathbf{y}_a = T_s\mathbf{A}_\tau^H\mathbf{r}$ and the DMF output vector $\mathbf{y}_d = T_s\mathbf{D}_\tau^H\mathbf{r}$ of components

$$y_a(k) = \int r(t)g(t - kT - \tau)\,dt = [r(t+\tau) * g(-t)]_{t=kT} \tag{86}$$

$$y_d(k) = \int r(t)g'(t - kT - \tau)\,dt = [r(t+\tau) * g'(-t)]_{t=kT} \tag{87}$$

respectively. The samples of $T_s\mathbf{A}_\tau^H\mathbf{r}$ and of $T_s\mathbf{D}_\tau^H\mathbf{r}$ correspond to stationary processes as both matched and derivative matched filter outputs are decimated at one sample per symbol. In addition, $\mathbf{R}_{aa} = E(T_s\mathbf{A}_\tau^H\mathbf{r})(T_s\mathbf{A}_\tau^H\mathbf{r})^H$ is the auto-correlation (Toeplitz) matrix of $\mathbf{y}_a$. We can now write, using Proposition B.3 in Appendix B, that

$$\begin{aligned}\frac{1}{2}\nabla_c^{as}(\tau) = \lim_{L\to\infty}\frac{1}{L}\,\mathrm{Re}[\mathrm{tr}((T_s\mathbf{A}_\tau^H\mathbf{A}_\tau)^{-1}\mathbf{R}_{ad} - (T_s\mathbf{A}_\tau^H\mathbf{A}_\tau)^{-1}\\ \cdot (T_s\mathbf{A}_\tau^H\mathbf{D}_\tau)(T_s\mathbf{A}_\tau^H\mathbf{A}_\tau)^{-1}\mathbf{R}_{aa})]\\ = \mathrm{Re}\int_{-1/2}^{1/2} S_{aa}^{-1}(e^{j2\pi F})(S_{ad}^{\mathbf{r}}(e^{j2\pi F})\\ - S_{aa}^{-1}(e^{j2\pi F})S_{ad}(e^{j2\pi F})S_{aa}^{\mathbf{r}}(e^{j2\pi F}))\,dF\end{aligned} \tag{88}$$

where $S_{ad}^{\mathbf{r}}(e^{j2\pi F})$ is the cross-spectrum between $y_a(k)$, which is the decimated MF, and $y_d(k)$, which is the decimated DMF output; in addition, $S_{aa}^{\mathbf{r}}(e^{j2\pi F})$ is the spectrum of $y_a(k)$. $S_{ad}^{\mathbf{r}}(e^{j2\pi F})$ and $S_{aa}^{\mathbf{r}}(e^{j2\pi F})$ are obtained from the Fourier transform of the central row of $\mathbf{R}_{ad}$ and $\mathbf{R}_{aa}$, respectively. $S_{aa}(e^{j2\pi F})$ and $S_{ad}(e^{j2\pi F})$ are previously defined in (27) and (29). We are now in a position to represent the asymptotic expression $(1/2)\nabla_c^{as}(\tau)$ in terms of filters.

Let us introduce the following intermediate property defined on two stationary processes $x_1(n)$ and $x_2(n)$, two filters $h_1(n)$ and $h_2(n)$, and their outputs $y_1(n) = x_1(n) * h_1(n)$ and $y_2(n) = x_2(n) * h_2(n)$

$$S_{y_1y_2}(e^{j2\pi F}) = H_1(e^{j2\pi F})H_2^*(e^{j2\pi F})S_{x_1x_2}(e^{j2\pi F}) \tag{89}$$

where $R_{x_1x_2}(k) \leftrightarrow S_{x_1x_2}(e^{j2\pi F})$ and $R_{y_1y_2}(k) \leftrightarrow S_{y_1y_2}(e^{j2\pi F})$ are the cross-correlation/cross-spectrum pairs between the two input and the two output processes, respectively.

Let us further define the following sequences and spectra

$$\begin{aligned}y_a'(k) &= y_a(k) * h_a(k), \qquad H_a(e^{j2\pi F}) = S_{aa}^{-1}(e^{j2\pi F})\\ y_d'(k) &= y_d(k) - h_{ad}(k) * y_a'(k)\\ H_{ad}(e^{j2\pi F}) &= S_{ad}(e^{j2\pi F}).\end{aligned} \tag{90}$$

Defining $y_{ad}'(k) = h_{ad}(k) * y_a'(k)$, the cross-correlation $R_{y_a'y_d'}(k-k') = E(y_a'(k)y_d'^*(k'))$ is

$$E(y_a'(k)y_d'^*(k')) = E(y_a'(k)y_d^*(k')) - E((y_{ad}'(k'))^*y_a'(k)). \tag{91}$$

Hence, using property (89) for both terms in (91), the corresponding cross-spectrum is

$$\begin{aligned}S_{y_a'y_d'}(e^{j2\pi F}) &= H_a(e^{j2\pi F})S_{y_ay_d}(e^{j2\pi F}) - H_{ad}^*(e^{j2\pi F})\\ &\quad\cdot S_{y_a'y_a'}(e^{j2\pi F}) = H_a(e^{j2\pi F})S_{y_ay_d}(e^{j2\pi F})\\ &\quad - H_{ad}^*(e^{j2\pi F})|H_a(e^{j2\pi F})|^2 S_{y_ay_a}(e^{j2\pi F})\end{aligned} \tag{92}$$

but $H_{ad}(e^{j2\pi F})$ and $H_a(e^{j2\pi F})$ are defined in (90), $S_{y_ay_a}(e^{j2\pi F}) = S_{aa}^{\mathbf{r}}(e^{j2\pi F})$, and $S_{y_ay_d}(e^{j2\pi F}) = S_{ad}^{\mathbf{r}}(e^{j2\pi F})$. Therefore, taking Re[·] in (92)

$$\begin{aligned}\mathrm{Re}S_{y_a'y_d'}(e^{j2\pi F}) &= \mathrm{Re}(S_{aa}^{-1}(e^{j2\pi F})\\ &\quad\cdot [(S_{ad}^{\mathbf{r}}(e^{j2\pi F}) - S_{aa}^{-1}(e^{j2\pi F})\\ &\quad\cdot S_{ad}(e^{j2\pi F})S_{aa}^{\mathbf{r}}(e^{j2\pi F}))])\end{aligned} \tag{93}$$

which shows that

$$\begin{aligned}\tfrac{1}{2}\nabla_c^{as}(\tau) &= \mathrm{Re}\int_{-1/2}^{1/2} S_{y_a'y_d'}(e^{j2\pi F})\,dF\\ &= \mathrm{Re}(E[y_a'(k)y_d'^*(k)])\end{aligned} \tag{94}$$

thus proving that the filters as defined in (90) synthesize the gradient $\nabla_c^{as}(\tau) = \lim_{L\to\infty}(T_s/L)\nabla_c(\tau) = E(\hat{\nabla}_c(\tau))$ asymptotically. The scheme is depicted in Fig. 4.

B. WMF and ODMF Analog Expressions

We can now translate the discrete filter expressions in (90) to equivalent filters in the analog domain. Equations (86) and (87) allow us to establish $y_a(k)$ and $y_d(k)$ as samples of the corresponding analog convolutions. We can define the analog equivalent filters $h_a^\delta(t)$ and $h_{ad}^\delta(t)$

$$h_a^\delta(t) = \sum_{k=-\infty}^{+\infty} h_a(k)\delta(t-kT),$$

$$H_a^\delta(f) = \sum_{k=-\infty}^{+\infty} h_a(k)e^{-j2\pi fkT}$$

$$h_{ad}^\delta(t) = \sum_{k=-\infty}^{+\infty} h_{ad}(k)\delta(t-kT)$$

$$H_{ad}^\delta(f) = \sum_{k=-\infty}^{+\infty} h_{ad}(k)e^{-j2\pi fkT}.$$

The spectrum of $h_a^\delta(t)$ and of $h_{ad}^\delta(t)$ is $H_a^\delta(f) = S_{aa}^{-1}(e^{j2\pi fT})$ and $H_{ad}^\delta(f) = S_{ad}(e^{j2\pi fT})$, respectively, as defined in (28). Therefore

$$\begin{aligned}y_a'(k) &= [r(t+\tau) * g_c(t)]_{t=kT}\\ g_c(-t) &= g(t) * h_a^\delta(t)\end{aligned} \tag{95}$$

$$\begin{aligned}y_d'(k) &= [r(t+\tau) * d_c(t)]_{t=kT}\\ d_c(-t) &= g'(t) - h_a^\delta(t) * h_{ad}^\delta(t) * g(t)\end{aligned} \tag{96}$$

with the analog WMF and ODMF defined as $g_c(t)$ and $d_c(t)$, respectively. As we were to prove, their transforms are given by the expressions

$$G_c(f) = G^*(f) S_{aa}^{-1}(e^{j2\pi fT})$$
$$D_c(f) = G^*(f)(2\pi f - S_{aa}^{-1}(e^{j2\pi fT}) S_{ad}^*(e^{j2\pi fT})).$$

Note that in the ISI-free case, (96) reduces to

$$d_c(-t) = g'(t) - \frac{1}{E_g}\sum_{n=-\infty}^{\infty} R_g'(nT) g(t-nT) \tag{97}$$

where $R_g'(t) = (d/dt)R_g(t)$, in agreement with [8].

C. Self-Noise Cancellation in the ODMF

Equation (88) relates the asymptotic gradient $\nabla_c^{as}(\tau)$ with the spectra at the MF and DMF outputs. For symmetry considerations, $\mathrm{Re}[S_{ad}(e^{j2\pi F})]$ is an odd function of the discrete frequency F as it is the cross-spectrum between the decimated MF and DMF filters [see (29)]. Hence, given that $S_{aa}(e^{j2\pi F})$ and $S_{aa}^{\mathbf{r}}(e^{j2\pi F})$ are even functions of F, we have

$$\mathrm{Re}\int_{-1/2}^{1/2} S_{aa}^{-2}(e^{j2\pi F}) S_{ad}(e^{j2\pi F}) S_{aa}^{\mathbf{r}}(e^{j2\pi F})\, dF = 0 \tag{98}$$

and hence, from substitution of (93) into (94), that

$$\begin{aligned}\tfrac{1}{2}\nabla_c^{as}(\tau) &= \mathrm{Re}\int_{-1/2}^{1/2} S_{aa}^{-1}(e^{j2\pi F}) S_{ad}^{\mathbf{r}}(e^{j2\pi F})\, dF \\ &= \mathrm{Re}(E[y_a'(k) y_d^*(k)]).\end{aligned} \tag{99}$$

This expression suggests an alternative way to define an asymptotic filter different from the proposed ODMF on the constraint that it synthesizes the same asymptotic gradient. Equation (99) shows that a WMF/DMF scheme is valid because it does not change the output cross-correlation (94) on which $\nabla_c^{as}(\tau)$ is based; from (98), we have that $\mathrm{Re}(E[y_a'(k) {y'}_{ad}^*(k)]) = 0$ in Fig. 4. That is, the output of $S_{ad}(e^{j2\pi F})$, $y_{ad}'(k)$ is not correlated with $y_a'(k)$. Therefore, a WMF/ODMF scheme is equivalent, asymptotically, to a WMF/DMF scheme ($L \to \infty$).

In addition, if any filter $H_{ad}(e^{j2\pi F})$ (92) with $\mathrm{Re}(H_{ad}(e^{j2\pi F}))$ and odd function of F is used in (94), an infinity of valid asymptotic filters can be defined that generate the same $\nabla_c^{as}(\tau)$. This explains the apparent contradiction of obtaining different asymptotic filters with low-SNR-UML and CML TEDs for ISI-free pulses, when, save constants, the WMF and DMF coincide.

APPENDIX E
EVALUATION OF $E[(\mathbf{v}^H\mathbf{\Omega}\mathbf{v})^2]$

A. Proposition E.1

$$E[(\mathbf{v}^H\mathbf{\Omega}\mathbf{v})^2] = \mathrm{tr}^2(\mathbf{\Omega}) + (K-1)\mathrm{tr}(\mathbf{\Omega}^2)$$

$$\mathbf{\Omega} = \mathbf{B}^H\mathbf{B} = \sum_{i=1}^{L}\lambda_i \mathbf{u}_i\mathbf{u}_i^H = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$$

for $\mathbf{v}$ normally distributed, $E[\mathbf{v}\mathbf{v}^H] = \mathbf{I}$, and $K=3$ or $K=2$, depending on whether the $y_i = \mathbf{u}_i^H\mathbf{v}$ are real or complex.

Proof: By the special structure of $\mathbf{\Omega}$

$$\begin{aligned}E[(\mathbf{v}^H\mathbf{\Omega}\mathbf{v})^2] &= E\left(\sum_{i=1}^{L}\lambda_i|\mathbf{u}_i^H\mathbf{v}|^2\right)^2 \\ &= \sum_{i=1}^{L}\sum_{j=1}^{L}\lambda_i\lambda_j E\{|y_i|^2|y_j|^2\}\end{aligned} \tag{100}$$

where $y_i = \mathbf{u}_i^H\mathbf{v}$ are the components of $\mathbf{y} = \mathbf{U}^H\mathbf{v}$. The covariance matrix of $\mathbf{y}$ is $E[\mathbf{y}\mathbf{y}^H] = \mathbf{U}^H(E[\mathbf{v}\mathbf{v}^H])\mathbf{U} = \mathbf{I}$. Therefore, the y_i are normally distributed and independent. We may write

$$E\{|y_i|^2|y_j|^2\} = \begin{cases} E\{|y_i|^2\}E\{|y_j|^2\}, & i \neq j \\ E\{|y_i|^4\}, & i = j\end{cases} \tag{101}$$

and (100) becomes

$$\begin{aligned}E[(\mathbf{v}^H\mathbf{\Omega}\mathbf{v})^2] = &\left(\sum_{i=1}^{L}\lambda_i E\{|y_i|^2\}\right)^2 \\ &+ \sum_{i=1}^{L}\lambda_i^2(E\{|y_i|^4\} - E^2\{|y_i|^2\}).\end{aligned} \tag{102}$$

Because the y_i are normally distributed, $E\{|y_i|^2\} = 1$ and $E\{|y_i|^4\} = K$. Hence

$$\begin{aligned}E[(\mathbf{v}^H\mathbf{\Omega}\mathbf{v})^2] &= \left(\sum_{i=1}^{L}\lambda_i\right)^2 + \sum_{i=1}^{L}\lambda_i^2(K-1) \\ &= \mathrm{tr}^2(\mathbf{\Omega}) + (K-1)\mathrm{tr}(\mathbf{\Omega}^2)\end{aligned} \tag{103}$$

from $\mathbf{\Omega}^2 = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H\mathbf{U}\mathbf{\Lambda}\mathbf{U}^H = \mathbf{U}\mathbf{\Lambda}^2\mathbf{U}^H$, as we wanted to prove.

B. Proposition E.2

For $\mathbf{\Omega}$ positive definite, $\mathbf{v}$ not Gaussian (with $\mathbf{x} = \mathbf{\Gamma}^{1/2}\mathbf{v}$ as in Proposition A.1) and $E[\mathbf{v}\mathbf{v}^H] = \mathbf{I}$

$$\begin{aligned}&\lim_{L\to\infty}\frac{1}{L^2}E[(\mathbf{v}^H\mathbf{\Omega}\mathbf{v})^2] \\ &\quad = \lim_{L\to\infty}\frac{1}{L^2}(\mathrm{tr}^2(\mathbf{\Omega}) + (K-1)\mathrm{tr}(\mathbf{\Omega}^2))\end{aligned} \tag{104}$$

if $\lim_{L\to\infty}(1/L)\mathrm{tr}(\mathbf{\Omega})$ exists and there exists a finite supremum bound $\lambda_{\sup}$ to the λ_is when $L \to \infty$.

Proof: When $\mathbf{v}$ is not Gaussian, a correction term must be added to (103) (Proposition E.1 is valid when $\mathbf{v}$ is normal) to regain the original general expression in (100)

$$\frac{1}{L^2}E[(\mathbf{v}^H\mathbf{\Omega}\mathbf{v})^2] = \frac{1}{L^2}\left(\mathrm{tr}^2(\mathbf{\Omega}) + (K-1)\mathrm{tr}(\mathbf{\Omega}^2) + \sum_{i=1}^{L}\sum_{j=1}^{L}\lambda_i\lambda_j\, d_{ij}\right)$$

with $d_{ij} = E\{|y_i|^2|y_j|^2\} - 1 - (K-1)\delta_{ij}$, and δ_{ij} is the Kronecker delta. In virtue of the central limit theorem, if the components of $\mathbf{x}$ are independent, variables $y_i = \mathbf{u}_i^H\mathbf{v} = (\mathbf{u}_i^H\mathbf{\Gamma}^{-1/2})\mathbf{x}$ become normally distributed as L goes to infinity. In general, there will be some degree of dependence among the symbols $\mathbf{x}$

introduced by the encoder. However, any practical channel encoder uses finite memory (which is strictly true in the case of block encoder and asymptotically true in the case of convolutional encoders); therefore, we can still assume that the elements of $\mathbf{x}$ are block independent, and the central limit theorem is valid anyway. Therefore, as the variables y_i and y_j are uncorrelated, and they become normally distributed for large L, we conclude that they become independent. Then, when $L \to \infty$, each $y_i = \mathbf{u}_i^H \mathbf{v}$ is asymptotically normal, and $\lim_{L\to\infty} d_{ij} = 0$ (y_i is also asymptotically normal when $\mathbf{\Omega} = \mathbf{I}$ because we can set $\mathbf{\Omega} = \mathbf{U}\mathbf{U}^H$, with $\mathbf{U}$ unitary). Hence, as $\lim_{L\to\infty}(1/L)\sum_{i=1}^{L}\lambda_i$ exists (is finite) and the λ_is are bounded,

$$\lim_{L\to\infty} \frac{1}{L^2}\sum_{i=1}^{L}\sum_{j=1}^{L}\lambda_i\lambda_j\, d_{ij} = \lim_{L\to\infty} \frac{1}{L^2}\sum_{i=1}^{L}\sum_{j=1}^{L}\lambda_i\lambda_j(\lim_{L\to\infty} d_{ij}) = 0$$

and (104) is true.

REFERENCES

[1] M. Moeneclaey, "A simple lower bound on the linearized performance of practical symbol synchronizers," *IEEE Trans. Commun.*, vol. COMM-31, pp. 1029–1032, Sept. 1983.

[2] ——, "A fundamental lower bound on the performance of practical joint carrier and bit synchronizer," *IEEE Trans. Commun.*, vol. COMM-31, pp. 1007–1012, Sept. 1984.

[3] A. D'Andrea, U. Mengali, and R. Reggiannini, "The modified Cramer–Rao bound and its application to synchronization problems," *IEEE Trans. Commun.*, vol. 42, pp. 1391–1399, Feb./Mar./Apr. 1994.

[4] S. Parkvall and E. Ström, "Parameter estimation and detection of DS-CDMA signals subject to multipath propagation," in *Proc. IEEE/IEE Workshop Signal Process. Multipath Environments*, Glasgow, U.K., 1995.

[5] U. Mengali and A. D'Andrea, *Synchronization Techniques for Digital Receivers*. New York: Plenum, 1997.

[6] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1783–1795, Oct. 1990.

[7] A. D'Andrea and M. Luise, "Optimization of symbol timing recovery for QAM data demodulators," *IEEE Trans. Commun.*, vol. 44, pp. 399–406, Mar. 1996.

[8] M. Moeneclaey, "A comparison of two types of symbol synchronizers for which self-noise is absent," *IEEE Trans. Commun.*, vol. COMM-31, pp. 329–334, Mar. 1983.

[9] J. Riba, G. Vázquez, and S. Calvo, "Conditional maximum likelihhod frequency estimation for offset modulations," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, vol. VI, Seattle, WA, May 1998, pp. 3425–3428.

[10] J. Riba and G. Vázquez, "Conditional maximum likelihhod timing recovery," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, vol. IV, Phoenix, AZ, Mar. 1999, pp. 1817–1820.

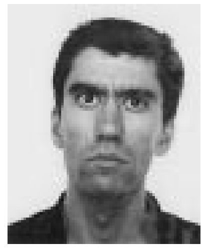
[11] G. Vázquez, J. Riba, G. B. Giannakis, Y. Hua, P. Stoica, and L. Tong, "Non-data-aided digital synchronization," in *Signal Process. Adv. Wireless Mobile Commun.*. Englewood Cliffs, NJ: Prentice-Hall, July 2000, vol. 2, Trends in Single- and Multi-User Systems, ch. 9.

[12] B. Ottersten, M. Viberg, and P. Stoica, "Exact and large sample maximum likelihood techniques for parameter estimation and detection," in *Radar Array Processing*. New York: Springer-Verlag, 1993, ch. 4.

[13] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood and Cramer–Rao bound," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 720–741, May 1989.

[14] B. Ottersten, M. Viberg, and T. Kailath, "Analysis of subspace fitting and ML techniques for parameter estimation from sensor array data," *IEEE Trans. Signal Processing*, vol. 40, pp. 590–600, Mar. 1992.

[15] M. Viberg, B. Ottersten, and T. Kailath, "Detection and estimation in sensor arrays using weighted subspace fitting," *IEEE Trans. Signal Processing*, vol. 39, pp. 2436–2449, Nov. 1991.



Jaume Riba was born in Barcelona, Spain, in 1966. He received the M.Sc. and Ph.D. degrees in telecommunications engineering from the Polytechnic University of Catalonia (UPC), Barcelona, in 1992 and 1997, respectively.

In 1992, he joined the Department of Signal Theory and Communications, UPC, as Assistant Professor and was promoted to Associate Professor in 1997. His current research interests are in the area of signal processing and communications, with particular emphasis on array processing and synchronization techniques. He has been involved in several signal processing research and development projects in the framework of the research programs of the European Space Agency.

Josep Sala was born in Callús, Spain, in 1967. He received the M.Sc. and Ph.D. degrees in telecommunications engineering from the Polytechnic University of Catalonia (UPC), Barcelona, Spain, in 1991 and 1995, respectively.

During 1992, he was working at the European Space Operations Centre (ESOC) of the European Space Agency (ESA), Darmstadt, Germany, in the area of software engineering for telemetry processing. From 1993 to late 1994, he held a grant from the Generalitat de Catalunya in support of the Ph.D. degree at the Department of Signal Theory and Communications, UPC. In 1994, he joined this department as Assistant Professor and was promoted to Associate Professor in 1997. His current research interests are in the area of signal processing and communications, with particular emphasis on array processing, adaptive filtering, and demodulation. He is also involved in the field of ASIC design for communications and signal processing applications in the context of departamental research and participation of projects with European industries.

Gregori Vázquez (SM'98) was born in Barcelona, Spain, in 1961. He received the M.Sc. and Ph.D. degrees in telecommunications engineering from the Polytechnic University of Catalonia (UPC), Barcelona, in 1984 and 1988, respectively.

He is Professor at the Department of Signal Theory and Communications, UPC. His general interests are signal processing and digital communications. He conducts research activities in digital synchronization and transmission techniques in time-varying frequency-selective channels. He has been involved in many research projects in wireless and satellite communications under the research programs of the European Union and the European Space Agency.

Dr. Vázquez is serving as Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING.