

A Nodata-Aided SNR Estimation Technique for Multilevel Modulations Exploiting Signal Cyclostationarity

Jaume Riba, *Senior Member, IEEE*, Javier Villares, *Member, IEEE*, and Gregori Vázquez, *Senior Member, IEEE*

Abstract—Signal-to-noise ratio (SNR) estimators of linear modulation schemes usually operate at one sample per symbol at the matched filter output. In this paper we propose a new method for estimating the SNR in the complex additive white Gaussian noise (AWGN) channel that operates directly on the oversampled cyclostationary signal at the matched filter input. Exploiting cyclostationarity proves to be advantageous due to the fact that a signal-free Euclidean noise subspace can be identified such that only second order moments of the received waveform need to be computed. The proposed method is nodata-aided (NDA), as well as constellation and phase independent, and only requires prior timing synchronization to fully exploit the cyclostationarity property. The estimator can also be applied to nonconstant modulus constellations without requiring any tuning, which is a feature not found in existing approaches. Implementation aspects and simpler suboptimal solutions are also provided.

Index Terms—Cyclostationarity, SNR estimation, second-order methods, spectral coherence, rate of innovation.

I. INTRODUCTION

ESTIMATION of the signal-to-noise ratio (SNR) is an important task in many digital communication receivers, since this parameter is a key indicator of link quality and can be used for optimal signal detection, information decoding, link adaptation, among others. In particular, a growing number of advanced transmission schemes such as diversity reception, adaptive coding and modulation (ACM), turbo decoding and turbo equalization, match their operation to the observed SNR, intending to push system performance to the achievable limits. For these reasons, estimating the SNR of digital communications signals is a timely research topic.

In those scenarios in which sufficiently long training sequences are available for SNR estimation, data-aided (DA) algorithms are recommended because they attain the ultimate performance given by the Cramer-Rao lower bound (CRLB) of the problem [1]. Otherwise, nodata-aided (NDA) *in-service*

SNR estimators operating on the unknown information-bearing signal must be adopted. In the NDA case, the numerical analysis of the CRLB in [2] concludes that both DA and NDA methods perform identically at high SNR (independently of the transmitted constellation) whereas NDA algorithms exhibit a serious performance loss for low and medium SNR values, especially in case of nonconstant modulus constellations.

An intensive review and comparison in the AWGN channel of the most prominent NDA SNR estimators in the literature is presented in [1] where the selection of the *best* alternative is shown to depend largely on the block length, the transmitted symbols alphabet, and the SNR range of interest. Among the NDA available techniques, the *second- and fourth-order moments* method (M_2M_4) (see [1], [3], and references therein) seems to be the one that exhibits the most robust performance for a wide range of SNR values and can be tailored to work with arbitrary constellations. However, its performance degrades dramatically at high SNR in case of nonconstant modulus constellations. Alternatively, near efficient SNR estimates can be obtained at high SNR using the *squared signal-to-noise variance* method (SNV) (see [1] and references therein), although this method is usually abandoned because it becomes excessively biased for medium-to-low SNR values. Regarding again the M_2M_4 estimator, Gao and Tepedelenlioglu [4] have shown that it belongs to a wider family of schemes that obtain the SNR as a function of ratios between moments. In this context, López-Valcarce and Mosquera [5] have considered the use of the signal sixth-order moment in order to extend the range of working SNR values in case of nonconstant modulus constellations.

In general, all the referred methods are intended to estimate the ratio of the discrete signal power to discrete noise power at the output of the matched filter evaluated at the optimal sampling instants. The signal and noise discrete sequences obtained in this manner are known to be wide-sense stationary random sequences. Thus, in order to estimate separately the power of each component (whose ratio is the desired SNR value in the case of complex channels), the moment-based methods presented before (e.g., M_2M_4) have to compute some *higher-order* sample moments of the noisy symbol-rate sequence, because the statistics become the unique difference between these two components. In doing so, the noise is assumed Gaussian and the statistics of the modulation becomes dependent on the constellation format.

A new approach is proposed in this paper for estimating the SNR in the complex AWGN channel. Instead of developing the

Manuscript received November 10, 2009; accepted June 27, 2010. Date of publication July 19, 2010; date of current version October 13, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Amir Leshem. This work was partially supported by the Spanish/Catalan Science and Technology Commission and by FEDER funds from the European Commission: TEC2007-68094-C02-02, CSD2008-00010, 2009SGR-1236.

The authors are with the Department of Signal Theory and Communications, Technical University of Catalonia (UPC), Barcelona, Spain (e-mail: jaume.riba@upc.edu; javier.villares@upc.edu; gregori.vazquez@upc.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSP.2010.2059017

SNR estimator at one sample per symbol, we start directly from the oversampled signal before matched filtering, although still assuming that the timing is perfectly known. It is noted that sampling the received signal at more than one sample per symbol (typically 2 to 4 samples per symbol) is not a serious practical concern, because this strategy is commonly used anyway in nowadays digital demodulators. Oversampling typically allows, for instance, a digital implementation of the matched filter, a digital interpolation for the correction of the timing error, a proper operation of digital timing error recovery schemes, as well as the implementation of fractionally-spaced equalizers. As a consequence of processing the oversampled data, the noise and the signal differ from the fact that the former is stationary while the latter is cyclostationary (apart from having different statistics). In fact, linear modulation schemes based on pulse amplitude modulation (PAM) belong to the class of signals that have a finite number of degrees of freedom per time-unit. This concept, also called *rate of innovation*, was introduced by Vetterli *et al.* in [6] and it is the core idea of the proposed estimator. As the rate of innovation of a PAM coincides with the symbol rate, and this is lower than its total spectral support, the useful signal is then confined to a specific subspace of the total space generated by the received signal. The identification of this subspace is what allows a more direct and robust noise and signal power estimation based only on second-order sample moments. The fact that the orthogonal signal subspace is affected only by the noise term leads to a SNR estimator that becomes unbiased at any SNR, even for nonconstant modulus constellations. Also, the resulting estimator is quadratic in the received data and hence not subject to the problematic use of higher order moments. Finally, the estimation algorithm is independent of the transmitted constellation and, therefore, it is a good candidate for implementing adaptive coding and modulation strategies [7].

In summary, the new SNR estimation technique proposed in this paper follows an alternative approach to previous methods: while the classical approach focuses on exploring the performance gain in using higher order statistics under the constraint of processing the signal at the symbol rate, the proposed approach is aimed at exploring the performance gain in processing the signal at the Nyquist rate under the constraint of using only second order statistics.

II. PROBLEM FORMULATION

Let $x(t)$ be a complex-valued linear modulation of the form

$$x(t) = \sum_k a[k]g(t - \varepsilon - kT) \quad (1)$$

with spectral support

$$I_x \triangleq \left(-\frac{1+\alpha}{2T}, \frac{1+\alpha}{2T} \right) \quad (2)$$

where $a[k]$ are the complex zero-mean, uncorrelated symbols, T is the symbol period, $g(t)$ is the unitary-energy pulse shape, and α is the roll-off factor that determines the excess of bandwidth with respect to the Nyquist bandwidth $1/T$.

Although the formulation of the problem can be applied to any pulse shape, for the sake of concreteness we will focus on the classical case of square-root-raised-cosine (SRRC) pulses, whose frequency response is given by

$$G(f) = \begin{cases} \sqrt{T} & 0 \leq |f| \leq \frac{1-\alpha}{2T} \\ \sqrt{T} \cos\left(\frac{\pi T}{2\alpha}\left(|f| - \frac{1-\alpha}{2T}\right)\right) & f \in I_w \\ 0 & |f| > \frac{1+\alpha}{2T} \end{cases} \quad (3)$$

and where I_w will denote the roll-off band support given by

$$I_w \triangleq \left(-\frac{1+\alpha}{2T}, -\frac{1-\alpha}{2T} \right) \cup \left(\frac{1-\alpha}{2T}, \frac{1+\alpha}{2T} \right). \quad (4)$$

Consider the received signal $y'(t) = x(t) + w'(t)$ where $w'(t)$ is wide-sense stationary, complex, circular, white noise whose power spectral density is $2N_0$ corresponding to the assumed AWGN channel model. This signal is passed through an antialiasing filter which is assumed to have flat (unitary) spectral response inside the spectral support of $x(t)$ and null spectral response at frequencies $|f| > N/(2T)$ (with $N \geq 2$). In this manner, the signal at its output can be written as $y(t) = x(t) + w(t)$ where $w(t)$ is finite-power complex circular noise. The processing of signal $y(t)$ developed in this paper admits both analog and digital implementations. In this last case, the integer N defined before is associated with the minimum number of samples per symbol that should be employed at the receiver to avoid both signal and noise aliasing. Then, $F_s = N/T$ will denote the sampling rate, and $T_s = T/N$ the sampling interval.

It should be noticed that no assumption is made on the transition bands of the antialiasing filter. As in these bands there is only noise contribution, side information about the transition bands of the antialiasing filter can trivially be used if available in order to improve the noise power estimation. However, these bands can be very narrow in practice and/or subject to unknown interference from adjacent channels. Therefore, in order to avoid bias on the noise power estimate, it becomes clear that the challenging problem should be focused on estimating the noise power by relying only on the information available inside the useful signal band. In addition, the constraint of processing only the useful signal band is also imposed in the design of the new estimator, in order to make a fair comparison with previous methods in the final simulation results, which represents a worst-case scenario for the proposed technique. Note that this distinction is not necessary for the classical SNR estimation schemes that work at the matched filter output, because the matched filter automatically suppresses (almost) all the out-of-band noise plus interference.

The main contribution of the paper is to show that the digital modulation signal $x(t)$ before matched filtering spans a subspace with a *rate of innovation* [6] lower than that of the additive noise, such that a subspace technique can be developed for the estimation of the in-band noise power as well as for the SNR, by processing directly the oversampled (or analog) signal before the matched filter assuming that the timing is known. In this manner, it is possible to obtain an estimate of the noise power \hat{N}_0 (the one affecting the received symbols) that is not biased by the signal power, by using *only* second order statistics, which

constitutes the main focus of the present contribution. Once this estimate is available, a SNR estimate $\hat{\rho}$ is obtained as follows:

$$\hat{\rho} = \frac{\hat{E}_T - \hat{N}_0}{\hat{N}_0}$$

where \hat{E}_T is an estimate of the total signal plus noise power that can also be easily obtained using *only* second order statistics from the observed symbols at the matched filter output. The problem is first addressed in the frequency domain, where its statistical analysis becomes simpler. Then, asynchronous and synchronous implementations in the time-domain are proposed, where they make use of the timing parameter in a different manner.

III. SIGNAL AND NOISE SUBSPACES INDUCED BY CYCLOSTATIONARITY

As the key signal property to be exploited is related with the cyclic spectral density of the signal, it is convenient to first formulate the problem in the frequency domain. To this end, let us consider the Fourier transform of a finite (arbitrarily large) segment of size MT of the complex envelope of the received signal $y(t)$:

$$Y(f) = \int_{-\frac{MT}{2}}^{\frac{MT}{2}} y(u) e^{-j2\pi fu} du = X(f) + W(f) \quad (5)$$

where $X(f)$ and $W(f)$ are the finite-size Fourier transforms of the signal and noise components, respectively.

It is well known [9] that cyclostationary signals exhibit spectral correlation at frequency values separated by the fundamental cycle frequency $1/T$. Motivated by this fact, let us define the frequency pair $\{\nu_+, \nu_-\}$ as

$$\begin{aligned} \nu_+ &= \nu + \frac{1}{2T} \\ \nu_- &= \nu - \frac{1}{2T} \end{aligned} \quad (6)$$

where the frequency variable ν is introduced to scan the roll-off band I_w defined in (4) and $\nu = 0$ refers to the center of the roll-off band at frequencies $\pm 1/2T$. Using these two frequency bins, we can now take the pair of random variables $Y(\nu_+)$ and $Y(\nu_-)$ and define the normalized two-dimensional frequency-domain observation vector as

$$\mathbf{o}_y(\nu) \triangleq \frac{1}{\sqrt{MT}} \begin{pmatrix} Y(\nu_+) \\ Y(\nu_-) \end{pmatrix} = \mathbf{o}_x(\nu) + \mathbf{o}_w(\nu) \quad (7)$$

where $\mathbf{o}_x(\nu)$ and $\mathbf{o}_w(\nu)$ are the signal and noise components of the frequency-domain observation vector, and they are implicitly defined in a similar manner. The normalization factor $1/\sqrt{MT}$ in the previous definition is convenient when the asymptotic case of $M \rightarrow \infty$ is studied, as done in the sequel.

As the signal and noise vector components are uncorrelated, we can express the second order statistics of the frequency-domain observation vector as follows:

$$\mathbf{S}_y(\nu) = E \{ \mathbf{o}_y(\nu) \mathbf{o}_y^H(\nu) \} = \mathbf{S}_x(\nu) + \mathbf{S}_w(\nu) \quad (8)$$

where the 2×2 matrices $\mathbf{S}_x(\nu)$ and $\mathbf{S}_w(\nu)$ are the spectral coherence matrices of the signal and noise components, respectively, and $\mathbf{S}_y(\nu)$ the spectral coherence matrix of the received signal. Note that all these matrices are defined for a specific value of the auxiliary frequency ν , from which a specific frequency pair $\{\nu_+, \nu_-\}$ becomes fixed through (6).

In order to understand the particular structure of the coherence matrix of the received signal, let us first focus on its signal component $\mathbf{S}_x(\nu)$. The key point is to realize that this matrix possesses a low-rank structure for PAM signals, and this aspect constitutes the basis of the proposed noise power estimation approach. To see this, let us take the expression of the finite-size Fourier transform of the process as defined in (5), and compute the (generic) correlation between random variables $X(\nu_1)$ and $X(\nu_2)$, that is the term $E\{X(\nu_1)X^*(\nu_2)\}$, which represents the frequency-domain coherence that the useful signal exhibits at frequencies ν_1 and ν_2 . After interchanging the expectation operator and the time-domain integrals, the previous term becomes

$$\begin{aligned} E\{X(\nu_1)X^*(\nu_2)\} &= \int_{-\frac{MT}{2}}^{\frac{MT}{2}} \int_{-\frac{MT}{2}}^{\frac{MT}{2}} R_x(u_1, u_2) e^{-j2\pi(\nu_1 u_1 - \nu_2 u_2)} du_1 du_2 \end{aligned}$$

where the statistical autocorrelation function of the process is defined in the standard manner as $R_x(u_1, u_2) = E\{x(u_1)x^*(u_2)\}$.

Making the substitutions $u_1 = t + \tau/2$ and $u_2 = t - \tau/2$, and considering the asymptotic case of $M \rightarrow \infty$, the previous double integral can be expressed as a single integral as follows:

$$\begin{aligned} \frac{1}{MT} E\{X(\nu_1)X^*(\nu_2)\} &\rightarrow \int_{-\infty}^{\infty} R_x^{\nu_1 - \nu_2}(\tau) e^{-j2\pi \frac{\nu_1 + \nu_2}{2} \tau} d\tau \\ &= S_x^{\nu_1 - \nu_2} \left(\frac{\nu_1 + \nu_2}{2} \right) \end{aligned} \quad (9)$$

where we have used that $\lim_{T_o \rightarrow \infty} (1/T_o) \int_{-T_o/2}^{T_o/2} R_x(t + \tau/2, t - \tau/2) e^{-j2\pi\beta t} dt = R_x^\beta(\tau)$ is the cyclic autocorrelation function, and its Fourier transform $S_x^\beta(f)$ is the cyclic spectrum of the digital modulation $x(t)$.¹

¹For a more in-depth study of the concepts of cyclic autocorrelation and cyclic spectrum, the reader is referred to [8] which gives a concise survey of the literature on cyclostationarity. On the other hand, for clarity reasons in the mathematical developments, it is worth noting that we are using the symmetric version of the cyclic autocorrelation function in our analysis, although the asymmetric one, given by $R_x^\beta(\tau)_{\text{asym}} = \lim_{T_o \rightarrow \infty} (1/T_o) \int_{-T_o/2}^{T_o/2} R_x(t + \tau, t) e^{-j2\pi\beta t} dt$, has also been widely used in the literature. In any case, they are easily linked by the following relationship: $R_x^\beta(\tau) = e^{-j\pi\beta\tau} R_x^\beta(\tau)_{\text{asym}}$

It is well known (see [9] for instance) that, as the digital modulation $x(t)$ is wide-sense cyclostationary with cycle frequencies that are multiples of the symbol rate, $S_x^\beta(f)$ may be different from zero for values of β equal to these cycle frequencies. For signals having no more than 100% excess bandwidth (that is, $B \leq 1/T$), the unique cycle frequencies (values of β) to consider are $-1/T$, 0 and $1/T$. With this consideration, we can use the previous general expression to write the entries of the asymptotic spectral coherence matrix $\mathbf{S}_x(\nu)$ defined in (8) as

$$\begin{aligned} \frac{1}{MT} E\{X(\nu_+)X^*(\nu_-)\} &\rightarrow S_x^{1/T}(\nu) \\ \frac{1}{MT} E\{X(\nu_-)X^*(\nu_+)\} &\rightarrow S_x^{-1/T}(\nu) \\ \frac{1}{MT} E\{|X(\nu_+)|^2\} &\rightarrow S_x(\nu_+) \\ \frac{1}{MT} E\{|X(\nu_-)|^2\} &\rightarrow S_x(\nu_-) \end{aligned} \quad (10)$$

where $S_x(f) \triangleq S_x^0(f)$ is the power spectral density of $x(t)$. Therefore, we have that

$$\mathbf{S}_x(\nu) = \begin{pmatrix} S_x(\nu_+) & S_x^{1/T}(\nu) \\ S_x^{-1/T}(\nu) & S_x(\nu_-) \end{pmatrix}. \quad (11)$$

Finally, recalling that the cyclic spectral density of a PAM signal [9] is given by

$$S_x^{q/T}(f) = \frac{\sigma_a^2}{T} G\left(f + \frac{q}{2T}\right) G^*\left(f - \frac{q}{2T}\right) e^{-j2\pi q\varepsilon/T} \quad (12)$$

the spectral coherence matrix adopts the following structure:

$$\begin{aligned} \mathbf{S}_x(\nu) &= \frac{\sigma_a^2}{T} \\ &\times \begin{pmatrix} G(\nu_+)G^*(\nu_+) & G(\nu_+)G^*(\nu_-)e^{-j2\pi\varepsilon/T} \\ G(\nu_-)G^*(\nu_+)e^{j2\pi\varepsilon/T} & G(\nu_-)G^*(\nu_-) \end{pmatrix} \end{aligned} \quad (13)$$

The previous matrix is clearly rank-one ($\det \mathbf{S}_x(\nu) = 0$)². The physical interpretation of this fact is that random variables $X(\nu_+)$ and $X(\nu_-)$ are fully correlated, that is, one is just a scaled version of the other, and in particular

$$X(\nu_-) = \frac{G(\nu_-)}{G(\nu_+)} e^{j2\pi\varepsilon/T} X(\nu_+) \quad (14)$$

for those values of ν_+ such that $G(\nu_+) \neq 0$, the mathematical consequence is that the spectral coherence matrix can be expressed as the following outer product:

$$\mathbf{S}_x(\nu) = \frac{\sigma_a^2}{T} I^2(\nu) \mathbf{g}(\nu) \mathbf{g}^H(\nu) \quad (15)$$

where

$$\begin{aligned} \mathbf{g}(\nu) &= \frac{1}{I(\nu)} \begin{pmatrix} G(\nu_+)e^{-j\pi\varepsilon/T} \\ G(\nu_-)e^{j\pi\varepsilon/T} \end{pmatrix} \\ I(\nu) &= \sqrt{|G(\nu_+)|^2 + |G(\nu_-)|^2}. \end{aligned} \quad (16)$$

²Although we have assumed uncorrelated symbols to provide this result, it is not difficult to show that the rank-deficient property of the coherence matrix is preserved even if the symbols are correlated.

In the previous expression we have defined function $I(\nu)$ that, in general, depends on ν . However, if the pulse shape fulfills the Nyquist criterion for ISI-free transmission (at the matched filter output samples), which is known to be

$$\frac{1}{T} \sum_{k=-\infty}^{\infty} \left| G\left(f + \frac{k}{T}\right) \right|^2 = 1 \quad \text{for } \forall f \quad (17)$$

then function $I(\nu)$ becomes a constant, that is $I(\nu) = \sqrt{T}$, for any value of ν such that $|\nu| \leq (1 - \alpha/2)/T$. In that case, the unitary vector defining the signal space spanned for each value of ν can be expressed as follows:

$$\mathbf{g}(\nu) = \frac{1}{\sqrt{T}} \begin{pmatrix} G(\nu_+)e^{-j\pi\varepsilon/T} \\ G(\nu_-)e^{j\pi\varepsilon/T} \end{pmatrix}. \quad (18)$$

From the previous general result in (15), it becomes clear that matrix $\mathbf{S}_x(\nu)$ is a rank-1 two-dimensional matrix such that the signal subspace is that spanned by the unitary vector $\mathbf{g}(\nu)$, which is dependent on the timing parameter ε . We can then define the orthogonal signal subspace as that spanned by the following vector $\mathbf{g}_\perp(\nu)$:

$$\mathbf{g}_\perp(\nu) = \frac{1}{\sqrt{T}} \begin{pmatrix} G^*(\nu_-)e^{-j\pi\varepsilon/T} \\ -G^*(\nu_+)e^{j\pi\varepsilon/T} \end{pmatrix} \quad (19)$$

which fulfills that $\mathbf{g}_\perp^H(\nu)\mathbf{g}(\nu) = 0$ and $\mathbf{g}_\perp^H(\nu)\mathbf{g}_\perp(\nu) = 1$ for $|\nu| \leq (1 - \alpha/2)/T$.

An important remark is noted at this point. As the noise subspace is dependent on the timing parameter ε , it becomes clear that the noise power estimator developed under this perspective will require either the knowledge of this parameter or a joint estimation along with it. In other words, the in-band noise becomes accessible for linear modulation schemes only if the cyclostationarity property is fully exploited. Note that this limitation is not only associated with the proposed scheme, but also shared with all SNR schemes proposed in the literature which operate at the matched filter output correctly (synchronously) sampled at one sample per symbol.

Considering now the noise term, we can write its spectral coherence matrix as

$$\mathbf{S}_w(\nu) = \begin{pmatrix} S_w(\nu_+) & 0 \\ 0 & S_w(\nu_-) \end{pmatrix} \quad (20)$$

where the zero entries are a consequence of the fact that the noise process is wide-sense stationary³. Now, in order to cancel out the signal component for obtaining a measure of the noise level only, we can apply the orthogonal vector $\mathbf{g}_\perp(\nu)$ to the frequency-domain observation vector. In this way, we obtain a filtering process as a linear combination of two frequency bins of the signal. In particular, the output of this orthogonal filter at the associated frequency bin ν can be expressed as

$$\begin{aligned} z(\nu) &= \mathbf{g}_\perp^H(\nu) \mathbf{o}_y(\nu) = \frac{1}{T\sqrt{M}} \\ &\times [G(\nu_-)Y(\nu_+)e^{j\pi\varepsilon/T} - G(\nu_+)Y(\nu_-)e^{-j\pi\varepsilon/T}] \end{aligned} \quad (21)$$

³Notice that the noise statistical distribution is irrelevant and, therefore, it could be non-Gaussian.

The previous equation constitutes the basis of the proposed approach. In particular, the previous scalar product becomes just a projection of the received signal at frequency bin ν onto the space orthogonal to the one spanned by the useful signal. As a result, this projection succeeds in canceling out the cyclostationary component of the received signal, while the stationary component (the noise) is the only one allowed to pass, thus giving direct access to the in-band noise whose power is to be measured in a second stage.

Assuming that the timing parameter ε is known, the mentioned signal cancellation can be simply shown by using (5) and the linear dependence (14) in (21) to get

$$z(\nu) = \frac{1}{T\sqrt{M}} \times [G(\nu_-)W(\nu_+)e^{j\pi\varepsilon/T} - G(\nu_+)W(\nu_-)e^{-j\pi\varepsilon/T}] \quad (22)$$

which means that for a given value of frequency ν , the random function $z(\nu)$ becomes simply a linear combination of two frequency bins (ν_+ and ν_-) of the noise process, as a result of canceling the signal component.

At this point, it is also important to focus on the adequate spectral support of ν in the previous filtering definition in (21). Taking into account the imposed constraint that the out-of-band noise should not be used in the estimation process for the reasons explained in the introduction, that is $\nu_-, \nu_+ \in I_x$, it becomes clear from (6) that the possible values of ν should verify that

$$-\frac{1+\alpha}{2T} \leq \left\{ \nu - \frac{1}{2T}, \nu + \frac{1}{2T} \right\} \leq \frac{1+\alpha}{2T} \quad (23)$$

which leads to the following allowable range for ν :

$$I_g \triangleq \left(-\frac{\alpha}{2T}, \frac{\alpha}{2T} \right) \quad (24)$$

whose overall spectral support is α/T .

As a result of the previous support definition on the values of ν , and observing that the orthogonal vector in (19) explicitly depends on ν_+ and ν_- , it becomes clear that the obtained filtering process has a passband structure such that only the roll-off bands of the signal are used for noise power estimation. In particular, the frequency bins of the noise, $W(\nu)$, that are observed and used in (22) are those contained just in the roll-off band frequency support (4).

Finally, if the frequency range is constrained in this way, we can write the received signal coherence matrix of the observed frequency vector as

$$\mathbf{S}_y(\nu) = \sigma_a^2 \mathbf{g}(\nu) \mathbf{g}^H(\nu) + 2N_0 \mathbf{I}. \quad (25)$$

To conclude this section, and for the sake of clarity, Fig. 1 summarizes the three spectral supports considered so far associated to the frequency domain, I_x , I_g , and I_w , which will play an important role in the mathematical development of the estimator.

IV. FREQUENCY-DOMAIN FORMULATION

As $z(\nu)$ is the output of an orthogonal filtering process, it becomes signal-free and influenced only by the additive noise

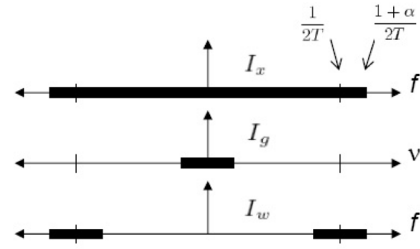


Fig. 1. Spectral supports considered in the study. I_x (2) is the spectral support of the digital modulation signal. I_g (24) is the spectral support of frequency variable ν associated to the orthogonal spectral vector $\mathbf{g}_\perp(\nu)$. I_w (4) is the spectral support of the noise that can be used to estimate its power without being affected by the signal (roll-off band support).

term $w(t)$. In fact, $|z(\nu)|^2/2$ constitutes an unbiased estimate of the (one-sided) noise density, because

$$\frac{1}{2} E|z(\nu)|^2 = \mathbf{g}_\perp^H(\nu) \mathbf{S}_y(\nu) \mathbf{g}_\perp(\nu) = N_0, \quad \text{for } \nu \in I_g \quad (26)$$

where the orthonormality property of $\mathbf{g}(\nu)$ and $\mathbf{g}_\perp(\nu)$ has been used ($\|\mathbf{g}_\perp(\nu)\|^2 = 1$ and $\mathbf{g}_\perp^H(\nu) \mathbf{g}(\nu) = 0$) in the previous equation. Note that the orthogonal signal subspace is the only one that can be used if an unbiased noise power estimate is desired, because estimating the noise term in the signal subspace would require the knowledge of the useful signal power, which is assumed unknown in most practical cases.

On the other hand, as $w(t)$ has normal distribution $z(\nu)$, which is a linear transformation of the noise process, possesses also a normal distribution. As a result, we can write its variance as

$$\text{VAR} \left(\frac{1}{2} |z(\nu)|^2 \right) = N_0^2, \quad \text{for } \nu \in I_g \quad (27)$$

where we have used again that $\mathbf{g}_\perp(\nu)$ is unitary. As it happens with the periodogram in spectral estimation theory, we can see that the estimator $|z(\nu)|^2/2$ is inconsistent, because its variance given by (27) does not decrease as M increases.

Considering that the useful spectral support is α/T and the fact that variables $z(\nu)$ and $z(\nu + 1/(MT))$ are uncorrelated for all ν due to noise stationarity and mutual bin orthogonality of the Fourier transform, we have that the total number of uncorrelated bins inside the support I_g is $\lfloor M\alpha \rfloor$, each one having identical variance. As a result, a (consistent) maximum likelihood estimate can be obtained as the sample mean of all available uncorrelated estimates

$$\hat{N}_0 = \frac{1}{2} \frac{1}{\lfloor M\alpha \rfloor} \sum_{l=-\lfloor M\alpha \rfloor/2}^{\lfloor M\alpha \rfloor/2-1} \left| z \left(\frac{l}{MT} \right) \right|^2. \quad (28)$$

In this manner, we get an estimator of N_0 whose variance exhibits now a consistent behavior

$$\text{VAR}(\hat{N}_0) = \frac{N_0^2}{\lfloor M\alpha \rfloor}. \quad (29)$$

At this point, it is important to observe that the proposed noise-power estimator has a variance that depends inversely on the roll-off factor of the digital modulation, in particular on the product $M\alpha$. In other words, for a given specified variance of the unbiased noise power estimate, modulations with small

roll-off will need more data with respect to modulations with high roll-off.

Finally, before going to the time-domain formulation, it is worth mentioning the existing link between our subspace formulation of the oversampled signal and the exploitation of the subspace structure proposed by Tong *et al.* in the well-known work [10], in which blind identification and equalization is made possible by using *only* second order statistics. It is seen in [10] that the structure of the time-domain covariance matrix is such that the signal space dimension is lower than the dimension of the observed oversampled signal, which means that it exists a noise subspace from which an unbiased noise power estimate can be obtained. Suitable methods exist for that purpose. For instance, a noise power estimate can be obtained by computing the singular value decomposition (SVD) of the covariance matrix, or by projecting the observed signal onto the noise eigenspace in the case that the eigenvectors for a given timing parameter are assumed to be known, for example, when there is no channel to be considered. Although this becomes a possible alternative for the derivation of our noise power estimator, the frequency-domain formulation proposed here has several advantages with respect to a direct time-domain approach:

- The computation of the SVD is not required. The reason is that, when examining the covariance structure at the frequency domain, it is seen that the frequency bins are correlated by pairs, which allows us to *only* consider multiple 2×2 coherence matrices instead of a single high dimension covariance matrix. These matrices have a trivial eigen-decomposition, such that the vector spanning the noise space can be trivially obtained from every 2×2 coherence matrix.
- The complexity of the problem does not substantially increase when the observed block size increases. This will become apparent in the next section where a time-domain estimator will be obtained by a direct translation from the frequency-domain formulation proposed here. In particular, it will be seen that the noise power can be estimated by simply computing the power at the output of a time-invariant filter, instead of directly projecting the observed data onto the noise eigenvectors of the entire covariance matrix. In fact, our (28) can be interpreted as an average of the estimated eigenvalues of the noise subspace, whose dimension increases as the block size increases, yielding as a result a consistent estimate.
- A closed-form expression for the variance of the resulting noise power estimate can be easily obtained in (29), which gives us a clear insight on the influence of the excess bandwidth of the signal on the estimator performance.

The time-domain formulation explored in the next section is obtained from a direct translation of the structure of the problem clearly identified here in the frequency-domain, yielding asynchronous baseband and synchronous passband alternatives for the implementation.

V. TIME-DOMAIN FORMULATION

The previous formulation in the frequency domain has allowed a simple mathematical path and a clear identification of the noise subspace. However, its direct implementation is not too much practical. In particular, $[M\alpha]$ points of the Fourier

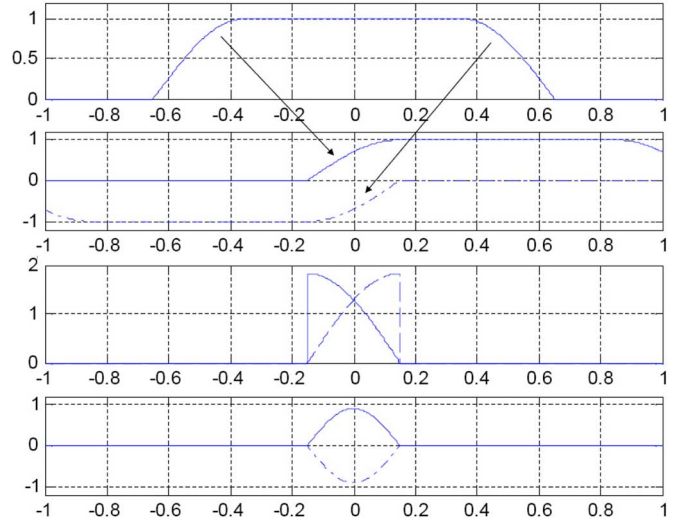


Fig. 2. (a) Frequency response of the original pulse. (b) Frequency shifts right and left an amount of $1/(2T)$ of the original signal. (c) Frequency response of filters $H_+(\nu)$ and $H_-(\nu)$. (d) Outputs of the previous filters where one can see the signal cancellation process.

transform in the roll-off bands of the signal need to be computed, with the added problem that M should be large enough in order to attain the mentioned asymptotic properties. In this section we propose a time-domain formulation that allows a simpler and sequential (instead of block) implementation, while still maintaining the assumed asymptotic properties.

For the derivation of a time-domain scheme, we will focus on the asymptotic case just by evaluating (28) when M goes to infinity

$$\lim_{M \rightarrow \infty} \hat{N}_0 = \frac{T}{2\alpha} \int_{-\frac{\alpha}{2T}}^{\frac{\alpha}{2T}} |z(\nu)|^2 d\nu. \quad (30)$$

This allows to make use of the Parseval's theorem (i.e., the unitary property of the Fourier transform) and compute the previous power in the time domain as follows:

$$\lim_{M \rightarrow \infty} \frac{T}{2\alpha} \int_{-MT/2}^{MT/2} |z(t)|^2 dt \quad (31)$$

where $z(t)$ is the inverse Fourier transform of $z(\nu)$. On the basis of the previous idea, we will next present an asynchronous baseband and a synchronous passband scheme implementing (31).

A. Asynchronous Baseband Scheme

In this first solution, a known timing error ε will be allowed in the processed data. In accordance with expression of $z(\nu)$ in (21), the derivation of $z(t)$ requires to process $y(t)$ as follows:

$$z(t) = e^{j\pi\varepsilon/T} (y(t)e^{-j\pi t/T}) * h_+(t) - e^{-j\pi\varepsilon/T} (y(t)e^{j\pi t/T}) * h_-(t) \quad (32)$$

where $h_+(t)$ and $h_-(t)$ are two *lowpass filters* with frequency response given by

$$\begin{aligned} H_+(\nu) &= \frac{1}{T\sqrt{M}} G(\nu_-) H_{\text{LPF}}(\nu) \\ H_-(\nu) &= \frac{1}{T\sqrt{M}} G(\nu_+) H_{\text{LPF}}(\nu) \end{aligned} \quad (33)$$

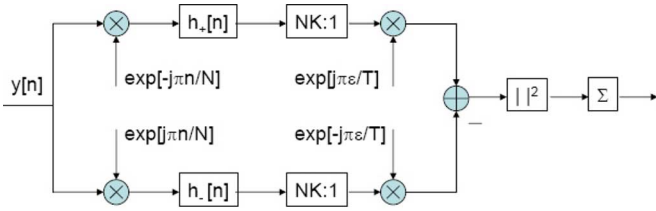


Fig. 3. Asynchronous baseband scheme for in-band noise power estimation. The subsampling is optional.

respectively, where $H_{\text{LPF}}(\nu)$ is an ideal *lowpass filter* with frequency support I_g , in accordance with the limits of the integral in (30). For SRRC pulses, the previous filters adopt the following frequency response:

$$\begin{aligned} H_+(\nu) &= \begin{cases} \frac{1}{\sqrt{MT}} \cos\left(\frac{\pi T f}{2\alpha} + \frac{\pi}{4}\right) & \text{for } \frac{-\alpha}{2T} \leq f \leq \frac{\alpha}{2T} \\ 0 & \text{otherwise} \end{cases} \\ H_-(\nu) &= H_+(-\nu) \end{aligned} \quad (34)$$

which is depicted in Fig. 2, along with the overall signal cancellation process. The direct time-domain implementation of the noise estimator is depicted in Fig. 3. It should be noted that this implementation requires the use of two complex-valued filters $h_+[n]$ and $h_-[n]$.

The two complex rotations $e^{j\pi\epsilon/T}$ and $e^{-j\pi\epsilon/T}$ after filters $H_+(\nu)$ and $H_-(\nu)$ introduce an angular rotation of their complex-valued outputs, which allows the signal component cancellation even if there is a timing error of ϵ seconds, with $|\epsilon| \leq T/2$. It is in these complex phase rotations where the timing information is used.

In addition, as the outputs of these two filters have a reduced spectral support of α/T , their outputs do not need to be sampled at the same sampling rate used for the received signal. In particular, it should be noted that a decimation factor of KN can be applied at that point, where $K \leq \lfloor 1/\alpha \rfloor$, which is an important consideration for real-time implementations.

B. Synchronous Passband Scheme

Another way of dealing with the timing error is to synchronize the received signal before starting the noise power estimation, typically acting on the sampling unit or performing a data interpolation according to the available timing error estimate. In this case, we can omit the complex rotations correcting the timing error in the previous scheme. When proceeding in this way, a simplification of the implementation can be derived, which makes use of a single real-valued passband filter and a synchronous subsampling at its output, as explained further on.

Let us define a *band-pass filter* $H(f)$ with the following frequency response:

$$H(f) = \begin{cases} -\frac{j}{T\sqrt{M}} G\left(f + \frac{1}{T}\right) & \text{for } -\frac{1+\alpha}{2T} \leq f \leq -\frac{1-\alpha}{2T} \\ \frac{j}{T\sqrt{M}} G\left(f - \frac{1}{T}\right) & \text{for } \frac{1-\alpha}{2T} \leq f \leq \frac{1+\alpha}{2T} \\ 0 & \text{otherwise} \end{cases} \quad (35)$$

where the (unitary) imaginary factor j is introduced in order to obtain a real-valued filter in the time-domain, without affecting at all the performance of the noise power estimator, as it is insensitive to signal phase. Thus, the main advantage of this

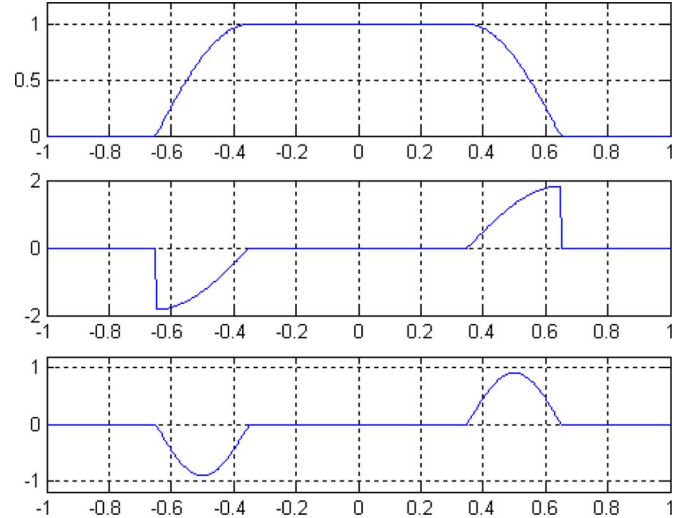


Fig. 4. (a) Frequency response of the matched filter. (b) Orthogonal passband matched filter. (c) Orthogonal passband matched filter output (down), for $\alpha = 0.3$.

scheme with respect to the asynchronous one proposed before, is that it requires a single real-valued filter instead of two complex-valued ones.

For SRRC pulses, the previous filter adopts the following frequency response:

$$H(f) = \begin{cases} \frac{j \operatorname{sign}(f)}{\sqrt{MT}} \sin\left(\frac{\pi T}{2\alpha} \left(|f| - \frac{1-\alpha}{2T}\right)\right) & \text{for } \frac{1-\alpha}{2T} \leq |f| \leq \frac{1+\alpha}{2T} \\ 0 & \text{otherwise} \end{cases} \quad (36)$$

which is depicted in Fig. 4, along with the frequency response of the overall pulse shape. This filter can be implemented either analogically or digitally, in this last case with a sampling rate fulfilling the Nyquist criterion.

Let $V(f) = Y(f)H(f)$ denote the Fourier Transform of the signal $v(t) = y(t) * h(t)$ at the output of this passband filter (see the third plot in Fig. 4). Notice that $V(f)$ contains the outputs of filters $h_+(t)$ and $h_-(t)$ in (32) but they are centered at frequencies $1/(2T)$ and $-1/(2T)$, respectively, instead of at frequency 0. Therefore, signal cancellation is still undone because it requires to subtract the content of $v(t)$ around frequencies $1/(2T)$ and $-1/(2T)$. To do so, we need to sample $v(t)$ by forcing spectral aliasing in such a way that aliasing performs subtraction of components at frequencies $1/(2T)$ and $-1/(2T)$ and, therefore, it leads to signal cancellation.

According to this, let us introduce the discrete Fourier transform of sequence $v(\epsilon + nT_s)$

$$V_d(e^{j2\pi f T_s}) = \frac{1}{T_s} \sum_{k=-\infty}^{\infty} V(f + k/T_s) e^{j2\pi(f+k/T_s)\epsilon} \quad (37)$$

which is defined in the interval $|f| < 1/(2T_s)$.

In order to allow signal cancellation it is simple to realize that the sampling period T_s should be an integer multiple of the symbol period, that is $T_s = KT$, where the decimation factor K can take any integer value $1, 2, \dots, \lfloor 1/\alpha \rfloor$. Notice that for odd values of K , cancellation (subtraction) occurs at discrete frequency $f = 0.5/K$ whereas for even values of K , it takes

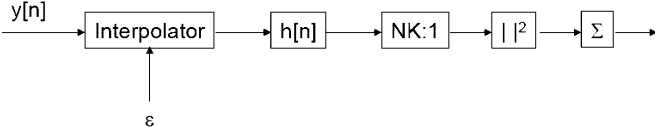


Fig. 5. Synchronous passband scheme for in-band noise power estimation. Subsampling at the symbol rate or below is now required.

place at discrete frequency $f = 0$, because there is then always a baseband alias. Notice also that while in the asynchronous baseband implementation (Section V-B) the decimation is optional, in the synchronous passband case it becomes mandatory. Moreover, it is also mandatory to perform this subsampling *synchronously* with the timing of the useful signal in order to effectively cancel-out the signal component.

If we now rewrite (30) in terms of (37), we obtain the following asymptotic result:

$$\begin{aligned} \lim_{M \rightarrow \infty} \hat{N}_0 &= \frac{T}{2\alpha} \int_{-\frac{1}{2T_s}}^{\frac{1}{2T_s}} \left| \sum_{k=-\infty}^{\infty} V(f + k/T_s) e^{j2\pi(f+k/T_s)\epsilon} \right|^2 df \\ &= \frac{TT_s^2}{2\alpha} \int_{-\frac{1}{2T_s}}^{\frac{1}{2T_s}} |V_d(e^{j2\pi f T_s})|^2 df \end{aligned} \quad (38)$$

that can be manipulated as follows:

$$\lim_{M \rightarrow \infty} \hat{N}_0 = \frac{TT_s}{2\alpha} \int_{-1/2}^{1/2} |V_d(e^{j2\pi F})|^2 dF$$

where the discrete frequency (F) has been normalized with respect to the sampling rate. If we invoke now the Parseval's theorem (i.e., the unitary property of the Fourier transform), we can express the result above in the time domain as follows:

$$\lim_{M \rightarrow \infty} \hat{N}_0 = \frac{T}{2\alpha} \lim_{M \rightarrow \infty} \frac{K}{M} \sum_{n=0}^{\frac{M}{K}-1} |v(\epsilon + nKT)|^2 \quad (39)$$

whose practical implementation is depicted in Fig. 5.

VI. SUBOPTIMAL IMPLEMENTATIONS

The orthogonal filtering proposed in (34) and (36) presents discontinuities in the frequency domain that pose difficulties on its implementation. In this section, we propose a suboptimal filter derived by introducing some degree of frequency smoothing at the filter edges, while still preserving the orthogonality (signal cancellation) property. We focus on the synchronous solution, but the concept can easily be applied in the asynchronous case as well. The modified filter is defined as

$$\tilde{H}(f) = H(f) \left(B\left(f - \frac{1}{2T}\right) + B\left(f + \frac{1}{2T}\right) \right) \quad (40)$$

where $B(f)$ is a real spectral window that fulfills the following properties:

$$\begin{aligned} B(\nu) &= 0 \quad \text{for } |\nu| \geq \frac{\alpha}{2T} \\ B(\nu) &\text{ is a continuous function} \\ B(\nu) &= B(-\nu) \\ \int_{-\frac{\alpha}{2T}}^{\frac{\alpha}{2T}} |B(\nu)|^2 d\nu &= \frac{\alpha}{T}. \end{aligned} \quad (41)$$

While the first and second properties are required to avoid sharp discontinuities of $\tilde{H}(f)$, the third one is required to preserve the signal cancellation property which, as explained, is a consequence of the mirror symmetry of function $V(f)$ around frequencies $\pm 1/(2T)$ (see bottom plot in Fig. 2). Finally, the fourth property is required to provide an unbiased noise power estimator.

In order to analyze the impact of this modification, we should note that it is equivalent to substitute the measure $z(\nu)$ in (21) by

$$\begin{aligned} \tilde{z}(\nu) &= \frac{1}{T\sqrt{M}} \times [G(\nu_-)B(\nu)Y(\nu_+)e^{j\pi\epsilon/T} \\ &\quad - G(\nu_+)B(-\nu)Y(\nu_-)e^{-j\pi\epsilon/T}]. \end{aligned} \quad (42)$$

As it can be trivially seen using (14), signal cancellation still occurs if $B(\nu) = B(-\nu)$ assuming correct timing synchronization. Then, (22) becomes

$$\begin{aligned} \tilde{z}(\nu) &= \frac{1}{T\sqrt{M}} \times [G(\nu_-)B(\nu)W(\nu_+)e^{j\pi\epsilon/T} \\ &\quad - G(\nu_+)B(-\nu)W(\nu_-)e^{-j\pi\epsilon/T}]. \end{aligned} \quad (43)$$

As a consequence, the mean and variance of $|\tilde{z}(\nu)|^2/2$ becomes spectrally shaped by the window in this manner

$$E\left(\frac{1}{2}|\tilde{z}(\nu)|^2\right) = N_0|B(\nu)|^2 \quad (44)$$

$$\text{VAR}\left(\frac{1}{2}|\tilde{z}(\nu)|^2\right) = N_0^2|B(\nu)|^4, \quad \text{for } \nu \in I_g. \quad (45)$$

The new estimator now becomes

$$\hat{N}_0 = \frac{1}{2} \frac{1}{[M\alpha]} \sum_{l=-\lceil[M\alpha]/2\rceil}^{\lceil[M\alpha]/2\rceil-1} \left| \tilde{z}\left(\frac{l}{MT}\right) \right|^2 \quad (46)$$

with a new statistical mean value for large M given by

$$\begin{aligned} E\{\hat{N}_0\} &= \frac{N_0}{[M\alpha]} \sum_{l=-\lceil[M\alpha]/2\rceil}^{\lceil[M\alpha]/2\rceil-1} \left| B\left(\frac{l}{MT}\right) \right|^2 \\ &\rightarrow N_0 \frac{\int_{-\frac{\alpha}{2T}}^{\frac{\alpha}{2T}} |B(\nu)|^2 d\nu}{\alpha/T} \end{aligned} \quad (47)$$

from which the fourth condition in (41) is confirmed. On the other hand, its variance and normalized variance (with respect

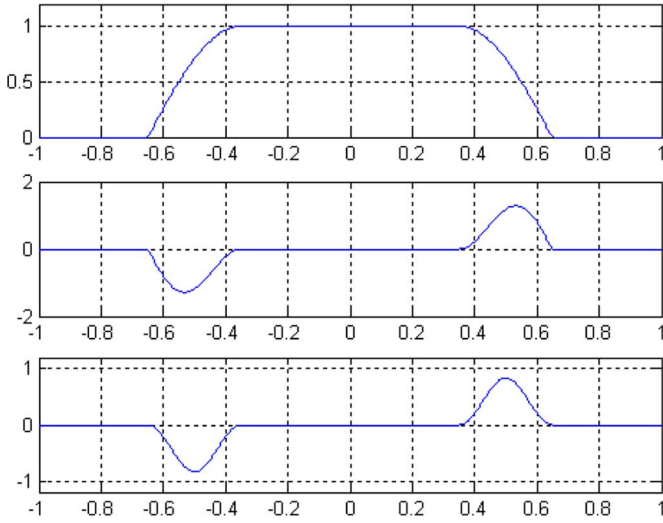


Fig. 6. Frequency response of the matched filter (up), modified orthogonal matched filter (center), and modified orthogonal matched filter output (down), for $\alpha = 0.3$. The required property to cancel-out the signal component is preserved, although the modified filter possesses less effective bandwidth than the true orthogonal one, which means that more data is required to achieve the same variance of the noise power density estimate. This is the tradeoff obtained in reducing the implementation complexity of the estimator. The modified orthogonal filter coincides with the orthogonal derivative matched filter derived within the context of conditional timing estimation [11].

to the squared true value of the parameter being estimated) are given by (for large M)

$$\begin{aligned} \text{VAR}(\hat{N}_0) &= \frac{N_0^2}{([\!M\alpha\!]^2)} \left[\sum_{l=-[\!M\alpha\!]/2}^{[\!M\alpha\!]/2-1} \left| B\left(\frac{l}{MT}\right) \right|^4 \right] \\ &\rightarrow \frac{N_0^2}{M\alpha} \gamma \\ \text{NVAR}(\hat{N}_0) &\rightarrow \frac{\gamma}{M\alpha} \end{aligned} \quad (48)$$

respectively, where it is assumed that $[\!M\alpha\!] \approx M\alpha$ for $M\alpha$ sufficiently large, and, $\gamma \geq 1$ is the variance amplification caused by the modification of the filter, which is given by

$$\gamma = \frac{\int_{-\frac{\alpha}{2T}}^{\frac{\alpha}{2T}} |B(\nu)|^4 d\nu}{\alpha/T}. \quad (49)$$

For the sake of concreteness, the following simple *ad hoc* design is proposed:

$$B(\nu) = \begin{cases} C \cos\left(\frac{\pi T \nu}{\alpha}\right), & \nu \in I_g \\ 0, & \text{otherwise} \end{cases} \quad (50)$$

By imposing the unbiasedness condition, we obtain that

$$\begin{aligned} C &= \sqrt{2} \\ \gamma &= \frac{3}{2} \end{aligned} \quad (51)$$

which means that the price of smoothing the filter edges using this window is an increase of 50% of the variance of noise

power spectral density estimation with respect to the optimal value given by (29). Interestingly, the filter $\tilde{H}(f)$ using the previous specific window coincides with the orthogonal derivative matched filter introduced in [11], where it was found in the context of conditional maximum likelihood timing estimation. In particular, $\tilde{H}(f)$ and $G(f)$ (the matched filter) are the two filters required to implement a self-noise free timing error detector. This can be especially convenient in practice because a single filter can be used for the joint timing and noise power estimation. The overall filtering process is depicted in Fig. 6.

VII. NOISE SUBSPACE ESTIMATOR (NSE) OF THE SNR

Estimating the SNR (ρ), which coincides with the E_s/N_0 for complex channels, requires an estimation of the symbol energy, E_s , along with the proposed estimation of N_0 . This estimate can be simply obtained from the sample energy at the matched filter output as follows⁴:

$$\begin{aligned} \hat{\rho} &= \frac{\hat{E}_s}{\hat{N}_0} = \frac{\hat{E}_T}{\hat{N}_0} - 1 \\ \hat{E}_T &= \frac{1}{M} \sum_{k=0}^{M-1} |r(kT)|^2 \end{aligned} \quad (52)$$

where $r(t) = y(t) * g^*(-t)$ is the matched filter output. The SNR estimator proposed in (52) will be called the NSE, which has the peculiarity of using solely second-order sample moments of the received signal. Note that \hat{E}_T is also a second-order sample moment that, along with \hat{N}_0 , contributes to the estimation errors of $\hat{\rho}$.

The normalized variance (NVAR) of the unbiased estimator $\hat{\rho}$ is defined as

$$\text{NVAR}(\hat{\rho}) = \frac{E\{(\hat{\rho} - \rho)^2\}}{\rho^2} \quad (53)$$

which coincides with the normalized mean squared error in the case of unbiased estimators, such as the proposed NSE.

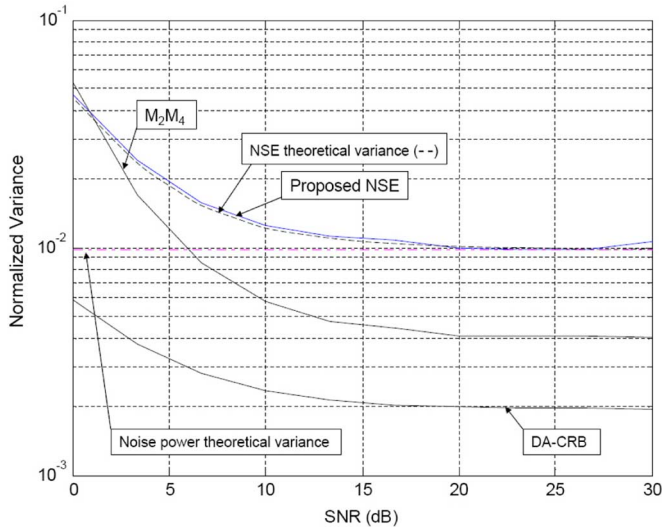
To evaluate the numerator of (53), we write $\hat{E}_T = E_T + \delta_E$ and $\hat{N}_0 = N_0 + \delta_N$, and make use of a first-order Taylor expansion of $\hat{\rho}$ around the point $\hat{\rho} = \rho$ to yield the following small error approximation:

$$\hat{\rho} \approx \rho - \frac{\rho + 1}{N_0} \delta_N + \frac{1}{N_0} \delta_E \quad (54)$$

Following an approach similar to the one in [5], NVAR is given by

$$\text{NVAR}(\hat{\rho}) = \frac{1}{E_s^2} E\{(\delta_E - (\rho + 1)\delta_N)^2\} \quad (55)$$

⁴It should be noted that the approach to estimate the SNR given in (52) will not work when co-channel interference, rather than thermal noise, is present because the interference may also be cyclostationary and it may introduce a positive or negative bias on the SNR estimate. In any case, it is also worth noting the SNR techniques based on higher order statistics will yield biased SNR estimates as well, because the interference may modify the statistics at the matched filter output in an unpredictable manner. Therefore, a further study would be required to make a comparative analysis of the impact of the co-channel interference on the referred SNR estimators in different, representative co-channel interference scenarios.

Fig. 7. NVAR of the SNR for PSK signals with roll-off $\alpha = 0.3$.

Then, assuming that δ_E and δ_N are uncorrelated, we have that

$$\begin{aligned} \text{NVAR}(\hat{\rho}) &= \frac{1}{E_s^2} E\{\delta_E^2\} + \frac{1}{N_0^2} \left(\frac{\rho+1}{\rho}\right)^2 E\{\delta_N^2\} \\ &= \frac{1}{E_s^2} E\{\delta_E^2\} + \frac{\gamma}{M\alpha} (1 + \rho^{-1})^2 \end{aligned} \quad (56)$$

where (29) has been recovered. Using closed form expressions of the even moments to expand $E\{\delta_E^2\}$ (see Appendix), we finally have

$$\text{NVAR}(\hat{\rho}) = \frac{(1 + \rho^{-1})^2 (1 + \gamma/\alpha) + c_4 - 2}{M} \quad (57)$$

where $c_p = E\{|a[k]|^p\}/E\{|a[k]|^2\}^p$ are the energy-normalized constellation moments.

It is seen then that the normalized SNR variance increases at low SNR values, and it is penalized for small roll-off factors. We can also see a persistent floor (at high SNR) inversely proportional to the roll-off factor and directly proportional to the kurtosis of the constellation given by the fourth-moment c_4 , which is equal to 1 for constant modulus constellations, and approaches to 2 for multilevel ones. In particular, for large SNR, we have that

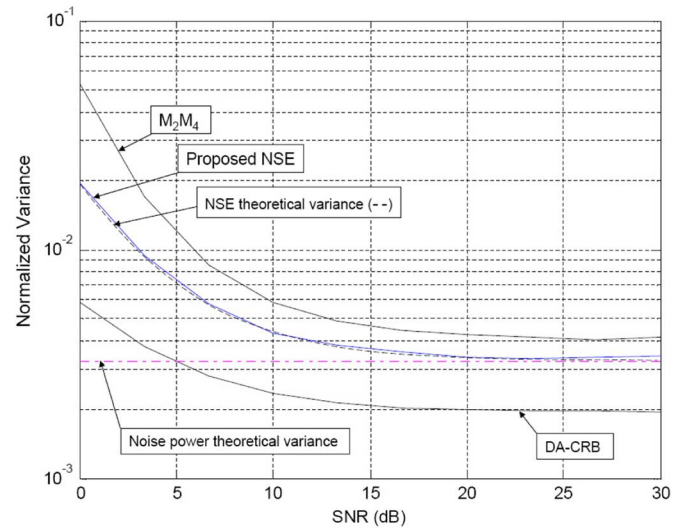
$$\text{NVAR}(\hat{\rho}) \rightarrow \frac{\gamma/\alpha + c_4 - 1}{M}. \quad (58)$$

Finally, a remarkable property is that, for constant modulus constellations and high SNR values, the NVAR floor coincides with the NVAR of the noise density derived in (48).

VIII. SIMULATION RESULTS

We show the results obtained via Monte Carlo simulation for a signal block length of $M = 512$ symbols, averaging 10^4 realizations for each SNR point, and using 201 coefficients in all FIR filter designs⁵. The synchronous passband scheme with the spectral window given by (50) has been used ($\gamma = 3/2$).

⁵The filter length has been selected large enough to avoid practical implementation losses.

Fig. 8. NVAR of the SNR for PSK signals with roll-off $\alpha = 0.9$.

In all simulated techniques, the received signal is sampled at two samples per symbol ($N = 2$). While the other techniques considered for comparison operate by using only a single sample per symbol at the matched filter output, the proposed technique uses all the available samples at the matched filter input for estimating the SNR. In both cases, however, identical signal and noise bandwidth are used to construct the resulting estimate, thanks to our constraint of using only the frequencies within the signal band when processing the input of the matched filter. This assumption is necessary in order to obtain a fair comparison with previous methods. Otherwise, the presence and utilization of the out-of-band noise would improve the performance of the new technique in estimating the noise power spectral level, as out-of-band noise would directly provide extra statistically independent information about the noise power spectral level that, besides, it is not contaminated by the useful signal.

Figs. 7 and 8 show the NVAR of SNR estimation for PSK signals (constant modulus, that is, $c_4 = 1$) with roll-off $\alpha = 0.3$ and $\alpha = 0.9$, respectively. The performance of the proposed NSE is shown along with that of the M_2M_4 . It is seen that for high values of α , the NSE outperforms the M_2M_4 while for low roll-off values, behavior is opposite. It can be seen how the NSE performance degrades as predicted in (57) when the roll-off decreases, while the behavior of M_2M_4 (and the CRB) is not affected by this parameter.

Fig. 9 shows the NVAR in the estimation of the SNR in the case of 16-APSK signals (nonconstant modulus), such as those specified in the DVB-S2 standard, consisting of an inner and outer ring with 4 and 12 symbols, respectively. The ring ratio is 2.57, the smallest value specified in the standard. It is seen that the performance of M_2M_4 degrades significantly, even tuning it with the known value of the constellation kurtosis (which is now $c_4 = 1.5432$). The performance of the sixth order estimator proposed in [5], which extends the usable range of the SNR, degrades as well⁶. On the contrary, the performance of the NSE only degrades slightly, as predicted in (57), without requiring

⁶In Fig. 9, the theoretical performance of the M_2M_4 and the sixth-order estimator is extracted from [5], where it was shown to be a tight lower bound of the actual variance of these estimators. Basically, both estimators' variance degrades as the SNR increases as predicted in [[5, eq. (15)]].

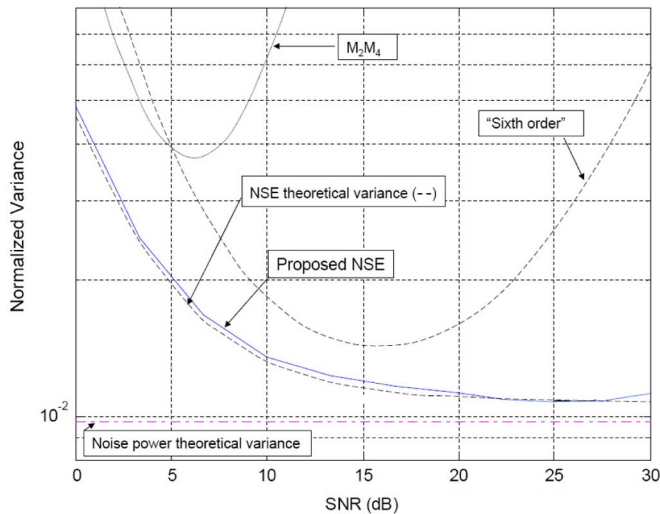


Fig. 9. NVAR of the SNR for 16-APSK signals with roll-off $\alpha = 0.3$ and ring ratio 2.57.

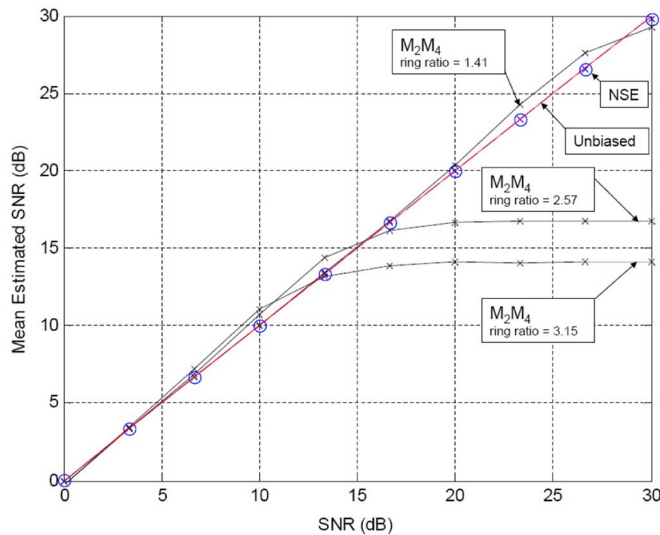


Fig. 10. Mean of the estimated SNR for 16-APSK signals of ring ratio 2.57 and 3.15, the extreme values specified in the DVB-S2 standard. The value 1.41 is also shown to highlight the sensitivity of moment-based methods to the constellation kurtosis, as opposed with the proposed second-order NSE approach.

any tuning. In fact, it is seen in Fig. 10 that the M_2M_4 becomes severely biased for $\text{SNR} > 15$ dB, while the NSE is always unbiased. In this figure, we have also added other values of the ring ratio of the 16-APSK modulation, in order to show how robust is the NSE method to the nonconstant modulus property of the digital modulations, compared with moment-based schemes.

It is worth noting, however, that for high SNR values, a SNR estimator that makes use of the demodulator decisions before the decoding stage can be used⁷. These techniques have the advantage of being asymptotically efficient for high SNR values, meaning that they achieve the CRB of the problem. At moderate SNR values, decision errors based on hard decisions would lead to an underestimation of the noise power (as the erroneous decision becomes closer to the received noisy symbol than the

⁷These SNR estimators are denoted by R_xDA in [1], and by *decision-directed* or *decision-based* approaches in other publications.

true one), which would produce a positive bias on the SNR estimate. This is an undesirable effect for instance in ACM scenarios where the SNR is continuously monitored and an action is expected to be taken when the estimated SNR decreases below a given threshold value. Additional drawbacks of DD techniques are that they completely fail at the lower range of SNR values, and that they require carrier phase synchronization as well as proper signal amplitude estimation (which is not required by the proposed technique) and the sensitivity to errors on these estimates becomes critical at high SNR values. Note that, in practice, it may be of interest to have an accurate estimate of the SNR even when this SNR is too low to implement DD estimation schemes. For example, LDPC codes in combination with 16-APSK modulation can provide a near error-free operation of SNR values around 8 dB, where the uncoded SER is about 0.35, which would provide completely unreliable SNR values when using a DD technique. As another example, ACM strategies may require accurate estimates of the SNR even at SNR ranges where the communication is not reliable because the link quality has been degraded, and feedback information is to be sent to the transmitter to properly modify the bit loading and coding rate policy. For all these reasons, the possibility of using a single SNR estimator for a wide range of SNR values is a desirable property that the proposed technique exhibits.

In general, it is seen that the performance of the new estimator is perfectly predicted by (57). The observed small mismatch at high SNR values is due to the limited number of coefficients used in the FIR filter designs, which leads to an imperfect signal cancellation at their output.

IX. CONCLUSION

A new NDA SNR estimator, called NSE, has been proposed by exploiting the full spectral coherence exhibited by PAM signals, which provides a noise subspace identification. Its performance is compared with the CRB and with the well-known moment-based M_2M_4 estimator. The main features of the NSE are the following:

- It operates on the received signal sampled at the Nyquist rate before matched filtering. Other NDA approaches, instead, are based on processing the signal at the matched filter output at the symbol rate. This is the peculiarity allowing a direct exploitation of the cyclostationarity property exhibited by the digital modulation, which provides a natural way of separating signal and noise which is not feasible if a symbol rate model is imposed from scratch.
- It does not require previous knowledge of the constellation format, in contrast with other moment-based approaches which need to be particularized for the constellation at hand (e.g., typically the M_2M_4 estimator needs to know the constellation kurtosis). Therefore, it becomes a good candidate in ACM schemes and/or in systems employing heterogeneous frames (see the work by Alvarez-Diaz *et al.* in [7] and references therein), so that the symbols in different subframes are drawn from different constellations.
- It performs well with nonconstant modulus constellations, while other approaches are known to completely fail at medium to high SNR values.

The weak point of the method is that the NVAR degrades for small values of the roll-off factor, requiring always some excess

bandwidth of the signal to properly work. This is a consequence of the fact that the estimator only uses sample second order moments of the signal, which is possible thanks to the exploitation of the cyclostationarity property. However, it is precisely this fact that provides its robustness compared with other approaches.

APPENDIX

The details for the derivation of (57) are provided. From (56), we only need to compute $E\{\delta_E^2\}$, that is the variance of \hat{E}_T in (52). As \hat{E}_T is the sum of M independent terms (symbol strobes plus noise), its variance is simply given by

$$E\{\delta_E^2\} = \frac{V}{M} \quad (59)$$

where V is the variance of the term $|r(kT)|^2$. Let us define M_n as the n th order moment of the strobe $r(kT)$. Using known closed-form expressions for the even-order moments (see [5, eqs. (3)–(4)], for instance), we can write

$$M_2 = E_s + N_0, \quad M_4 = c_4 E_s^2 + 4E_s N_0 + 2N_0^2. \quad (60)$$

We can now express V as $V = M_4 - M_2^2$ to yield

$$V = (c_4 - 1)E_s^2 + 2E_s N_0 + N_0^2 \quad (61)$$

and, therefore, the term $E\{\delta_E^2\}/E_s^2$ in (56) can be expressed as

$$\frac{1}{E_s^2} E\{\delta_E^2\} = c_4 - 1 + 2\rho^{-1} + \rho^{-2} = c_4 - 2 + (1 + \rho^{-1})^2. \quad (62)$$

Finally, using the previous expression in (56), we easily obtain (57).

REFERENCES

- [1] D. R. Pauluzzi and N. C. Beaulieu, "A comparison of SNR estimation techniques for the AWGN channel," *IEEE Trans. Commun.*, vol. 48, no. 10, pp. 1681–1691, Oct. 2000.
- [2] W. Gappmair, "Cramer-Rao lower bound for non-data-aided SNR estimation of linear modulation schemes," *IEEE Trans. Commun.*, vol. 56, no. 5, pp. 689–693, May 2008.
- [3] W. Gappmair, R. Lopez Valcarce, and C. Mosquera, "ML and EM algorithm for nondata-aided SNR estimation of linearly modulated signals," in *Proc. IEEE 6th Int. Symp. Commun. Syst., Netw. Digit. Signal Process. (CNSDSP 2008)*, Graz, Austria, Jul. 2008.
- [4] P. Gao and C. Tepedelenlioglu, "SNR estimation for nonconstant modulus constellations," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 865–870, Mar. 2005.
- [5] R. Lopez-Valcarce and C. Mosquera, "Sixth-order statistics-based non-data-aided SNR estimation," *IEEE Commun. Lett.*, vol. 11, no. 4, pp. 351–353, Apr. 2007.
- [6] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1417–1428, Jun. 2002.
- [7] M. Álvarez-Díaz, R. López-Valcarce, and C. Mosquera, "SNR estimation with heterogeneous frames," in *Proc. 4th Adv. Satellite Mobile Syst. Conf. (ASMS 2008)*, Bologna, Italy, Aug. 2008.
- [8] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal Process.*, vol. 86, no. 4, pp. 639–697, Apr. 2006.
- [9] W. A. Gardner, *Introduction to Random Processes, With Applications to Signals and Systems*, 2nd ed. New York: McGraw-Hill, 1990.
- [10] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inf. Theory*, vol. 40, no. 2, pp. 340–349, Mar. 1994.
- [11] J. Riba, J. Sala, and G. Vázquez, "Conditional maximum likelihood timing recovery: Estimators and bounds," *IEEE Trans. Signal Process.*, vol. 49, no. 4, pp. 835–850, Apr. 2001.



Jaume Riba (SM'05) was born in Barcelona, Spain, in 1966. He received the M.Sc. and Ph.D. degrees in telecommunications engineering from the Technical University of Catalonia (UPC), Barcelona, in 1992 and 1997, respectively.

In 1992, he joined the Department of Signal Theory and Communications (TSC) and the Signal Processing and Communications Group (SPC), UPC, as an Assistant Professor and was promoted to Associate Professor in 1997. His current research interests are in the area of signal processing, commu-

nications, and wireless location, with particular emphasis on array processing and digital synchronization techniques. He has been involved in several signal processing research and development projects with European Space Agency research programs. He was the Principal Investigator of the IST project European Mobile Integrated Location System (EMILY) at UPC (2001–2004).

Dr. Riba was a Guest Editor for the *IEEE Signal Processing Magazine* Special Issue on Signal Processing for Positioning and Navigation with Applications to Communications (July 2005). He has served as Associate Editor of the *IEEE TRANSACTIONS ON SIGNAL PROCESSING* in 2003 and is a Member of the Editorial Board and Associated Editor of the *International Journal on Navigation and Observation* (IJNO) since November 2006. He is a recipient of the 2003 Best Paper Award from the IEEE Signal Processing Society.



Javier Villares (S'01–M'05) was born in Barcelona, Spain, in 1974. He received the M.S. and Ph.D. degrees in telecommunications engineering from the Technical University of Catalonia (UPC), Barcelona, in 1999 and 2005, respectively.

He is a Professor with the Department of Signal Theory and Communications, UPC, since 2003. His general interests are digital communications, statistical signal processing, and information theory.



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 Technical University of Catalonia (UPC), Barcelona, in 1984 and 1988, respectively.

He is a Professor with the Department of Signal Theory and Communications, UPC. His general interests are statistical signal processing and digital communications.

Dr. Vázquez is a recipient of the 2003 Best Paper Award of the IEEE Signal Processing Society. He served as a member of the Editorial Board of the *IEEE SIGNAL PROCESSING MAGAZINE*, a member of the Signal Processing for Communications Technical Committee, and has also served as Associate Editor of the *IEEE TRANSACTIONS ON SIGNAL PROCESSING* from 1999 to 2005.