SNR Estimation for Multilevel Constellations Using Higher-Order Moments

Marcos Álvarez-Díaz, *Student Member, IEEE*, Roberto López-Valcarce, *Member, IEEE*, and Carlos Mosquera, *Member, IEEE*

***Abstract*—The performance of existing moments-based non-data-aided (NDA) estimators of signal-to-noise ratio (SNR) in digital communication systems substantially degrades with multilevel constellations. We propose a novel moments-based approach that is amenable to practical implementation and significantly improves on previous estimators of this class. This approach is based on a linear combination of ratios of certain even-order moments, which allow the derivation of NDA SNR estimators without requiring memory-costly lookup tables. The weights of the linear combination can be tuned according to the constellation and the SNR operation range. As particular case we develop an eighth-order statistics (EOS)-based estimator, showing in detail the statistical analysis that leads to the weight optimization procedure. The EOS-based estimators yield improved performance for multilevel constellations, especially for those with two and three amplitude levels. Monte Carlo simulations validate the new approach in a wide SNR range.**

***Index Terms*—Blind estimation, higher-order statistics, moments-based estimation, signal-to-noise ratio (SNR) estimation.**

I. Introduction

ESTIMATION of the signal-to-noise ratio (SNR) at the receiver side is an important task in existing and emerging communication systems, as many of them require SNR knowledge in order to achieve different goals, such as power control or adaptive coding and modulation [1], [2], and soft decoding [3]. Whereas SNR estimation is a relatively easy task with simple modulation formats such as quadrature phase-shift keying (QPSK), the increasing complexity of the modulation schemes used in modern systems poses significant challenges to current methods.

Existing SNR estimators can be classified according to a number of criteria. Data-aided (DA) estimators can be used when the receiver has knowledge of the transmitted symbols, in contrast to non-data-aided (NDA) estimators, which do not require such knowledge. DA estimators can be used in decision-directed (DD) mode by substituting the true transmitted symbols by the outputs of the decoder. Under a different classification, I/Q-based estimators make use of both the in-phase and quadrature components of the received signal, and thus require coherent detection; in contrast, envelope-based (EVB) estimators only make use of the received signal magnitude, and thus can be applied even if the carrier phase has not been completely acquired. This is important in applications in which the SNR must be estimated even if its value is so low as to preclude accurate synchronization and decoding. Concerning the sampling rate of the received signal, most estimators operate on baud-sampled data, although several estimators for oversampled data are also available [4], [5]. Most approaches focus on the single-input single-output (SISO) channel with additive white Gaussian noise (AWGN) and (quasi)static flat fading [4], [6]–[16], although [17] and [18] address the static frequency selective channel and time-varying flat fading channel cases respectively. SNR estimators for multiantenna receivers have also been recently proposed in [19], [20].

This paper addresses the SNR estimation problem under the AWGN SISO channel model. This applies either to single carrier systems or to multicarrier systems in which the SNR is to be estimated at each carrier. The *maximum likelihood* (ML) approach has been previously applied in this context, for both constant modulus (CM) [4], [6], [7], [11], [18], [21] and multilevel constellations [12] (I/Q-based), [13] (EVB). Although ML estimators provide good statistical performance, they tend to be computationally intensive. This motivates the development of simpler, suboptimal approaches such as those based on the moments of the envelope of the received signal, here referred to as M_n (n is the order). These methods belong to the class of NDA EVB estimators, requiring neither accurate carrier recovery, nor knowledge of the transmitted symbols. This flexibility, together with implementation simplicity, makes these estimators attractive for practical applications.

Early examples of moments-based methods are two estimators originally proposed for CM constellations [8], [9] built respectively upon the first- and second-order moments $(\mathrm{M}_1\mathrm{M}_2)$, and the second- and fourth-order moments $(\mathrm{M}_2\mathrm{M}_4)$. More recently, [10] proposed a family of estimators for multilevel constellations based on pairs of moments, containing $\mathrm{M}_1\mathrm{M}_2$ and $\mathrm{M}_2\mathrm{M}_4$ as particular cases. As it turns out, the performance of $\mathrm{M}_1\mathrm{M}_2$ and $\mathrm{M}_2\mathrm{M}_4$ is close to the Cramer–Rao Bound (CRB) for CM constellations [4], [8]. However, this is not so for multilevel constellations, for which the estimation variance considerably departs from the CRB as the SNR increases [10].

Recent efforts have been made in order to improve the quality of moments-based SNR estimators in these challenging settings.

Manuscript received July 06, 2009; accepted October 15, 2009. First published November 06, 2009; current version published February 10, 2010. This work was supported by the European Commission through the IST Programme under Contract IST-027393 SatNEx-II and by the Spanish Government under projects SPROACTIVE (ref. TEC2007-68094-C02-01/TCM) and COMONSENS (CONSOLIDER-INGENIO 2010 CSD2008-00010). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Alfred Hanssen.

The authors are with the Department of Signal Theory and Communications, University of Vigo, 36310 Vigo, Spain (e-mail: malvarez@gts.tsc.uvigo.es; valcarce@gts.tsc.uvigo.es; mosquera@gts.tsc.uvigo.es).

Digital Object Identifier 10.1109/TSP.2009.2036069

In [15], the observations are partitioned in different subsets corresponding to symbols of equal modulus, and then $\mathrm{M_2M_4}$ is applied to one of these subsets. Although the resulting performance is good for sufficiently high SNR, it degrades substantially for low SNR due to errors in the partition step. Alternatively, a sixth-order statistics-based estimator (henceforth referred to as $\mathrm{M_6}$) was proposed in [14] which makes use of M_2, M_4 and M_6, with a significant performance improvement over $\mathrm{M_2M_4}$ for multilevel constellations at intermediate and high SNR. In addition, a scalar parameter allows to tune the estimator to any particular constellation. The use of higher-order statistics has also been considered in other related contexts, such as modulation classification (see [22], [23] and the survey in [24]).

The results from [14] suggest that further improvements could be expected if higher-order moments are allowed in the construction of the estimator. With this in mind, this paper investigates general higher-order approaches for the moments-based SNR estimation problem. The main contributions are as follows.

- In Section II, we identify a novel family of estimators based on quotients of even-order moments which features an interesting property: all estimators in this family can be implemented via direct or iterative computations, thus avoiding the need for lookup tables.
- In Section III, we propose a general estimator built upon a linear combination of members of the previous family, which generalizes the sixth-order approach of [14] to any (even) order. The weights of the linear combination can be tuned to each particular constellation and nominal SNR operation range.
- To illustrate the potential of this method, in Section IV, we develop a statistical analysis of the particular case in which moments of up to eighth-order are allowed, whereas Section V presents two criteria for weight optimization. These approaches can be readily generalized to higher-order moments.
- As a result, we provide an eighth-order estimator with rather competitive performance for non-CM constellations, especially for two- and three-level constellations, as shown in Section VI.

II. Moments-Based SNR Estimation

A. Signal Model

We assume a quasistatic flat-fading channel model, with the symbol-rate samples at the matched filter output given by

$$r_k = \sqrt{S}x_k + n_k, \quad k = 1, \ldots, K \tag{1}$$

where x_k are the complex-valued transmitted symbols, $\sqrt{S}$ is the unknown channel gain, and n_k are complex circular i.i.d. Gaussian noise samples with unknown variance N. The symbols are i.i.d., drawn from a constellation which is known to the receiver and has I different amplitudes, being R_i and P_i, respectively, the ith amplitude and its associated probability ($i = 1, \ldots, I$). The *constellation moments* are denoted $c_p \doteq \mathrm{E}\{|x_k|^p\} = \sum_{i=1}^{I} P_i R_i^p$. Without loss of generality, an energy-normalized constellation is assumed, i.e., $c_2 = 1$.

Moments-based SNR estimation stems from the approach known as *method of moments* [25]. In our case, the goal is to find a function of the *sample moments* of the envelope

$$\hat{M}_p \doteq \frac{1}{K}\sum_{k=1}^{K} |r_k|^p \tag{2}$$

from which an estimate of the SNR $\rho \doteq S/N$ or the *normalized SNR* $z \doteq \rho/(1+\rho)$ can be derived.[1] $\hat{M}_p$ is a consistent, unbiased estimator of the *true moment* of the envelope, given by[2] [10], [26]

$$\begin{aligned} M_p &\doteq \mathrm{E}\left\{|r_k|^p\right\} \\ &= N^{\frac{p}{2}} \sum_{i=1}^{I} P_i \Gamma\left(\frac{p}{2}+1\right) e^{-\rho R_i^2} {}_1F_1\left(\frac{p}{2}+1; 1; \rho R_i^2\right) \end{aligned} \tag{3}$$

with $\Gamma(\cdot)$ the gamma function and ${}_1F_1(\cdot;\cdot;\cdot)$ the confluent hypergeometric function. For p even, (3) can be seen to admit a simpler form which depends only on the even constellation moments up to order p and the two unknowns ρ, N [14], [27]

$$M_{2n} = N^n \sum_{m=0}^{n} \frac{(n!)^2}{(n-m)!(m!)^2} c_{2m}\rho^m. \tag{4}$$

Note that (4) is a polynomial in ρ, and that in (3) and (4) the noise power N appears as a multiplicative factor. This latter fact allows to obtain pure functions of ρ through quotients of moments in which N vanishes. For example, [10] proposes the family of functions

$$f_{k,l}(\rho) \doteq \frac{M_k^l}{M_l^k}, \quad \text{for } k \neq l. \tag{5}$$

An SNR estimate can be then obtained from the sample moments by inversion of $f_{k,l}$

$$\hat{\rho}_{k,l} = f_{k,l}^{-1}\left(\frac{\hat{M}_k^l}{\hat{M}_l^k}\right). \tag{6}$$

In [10], Gao and Tepedelenlioğlu state that $f_{k,l}$ in (5) is monotonic.[3] This property is a requirement for any function $f(\rho)$ from whose inverse an estimator is to be derived. The $\mathrm{M_1M_2}$ and $\mathrm{M_2M_4}$ estimators are obtained for $(k,l) = (2,1)$ and (4,2), respectively.

B. New Family of NDA Moments-Based Estimators

The analytical inversion of $f_{k,l}$ is intractable in general, as one may expect from the complexity of (3) (an exception is the $\mathrm{M_2M_4}$ estimate, which admits a closed-form expression). In such cases one must resort to precomputed lookup tables (LUTs), as in [8], [10] for the $\mathrm{M_1M_2}$ estimator. Although LUTs are computationally efficient, their storage requirements,

[1] $z \in (0,1)$ is a one-to-one transformation of $\rho \in (0,\infty)$, so that for estimation purposes it is equivalent to derive an estimator of z and then undo the transformation to find ρ.

[2] There is typographical error in [10, Eqs. (3) and (4)]. The complex noise power is defined therein as $N = 2\sigma^2$, but the factor 2 does not appear in the expression of the moments.

[3] Although no proof is provided in [10], monotonicity holds for all constellations and all pairs (k,l) we have tested. Moreover, we have observed that $f_{k,l}$ is monotone decreasing when $k > l$, and increasing when $k < l$.

which are proportional to the desired accuracy, become an issue in hardware implementations with small memory space. This motivates the search for SNR estimators which do not rely on LUTs.

Consider the following family of functions built upon even-order moments:

$$f(\rho) \doteq \left(\prod_{i=1}^{U} M_{2n_i}^{p_i}\right) \Big/ M_2^Q, \quad n_i, p_i \in \mathbb{N} \tag{7}$$

where $Q = \sum_{i=1}^{U} n_i p_i$ in order to ensure that the quotient (7) does not depend on N. We refer to $2Q$ as the *statistical order* of (7). We assume that M_2 does not appear in the numerator (i.e., $n_i > 1$ for all i). Note that (7) can be written in terms of the functions $f_{k,l}$ from (6) as $f(\rho) = \prod_{i=1}^{U} f_{2n_i,2}^{p_i/2}(\rho)$, so it inherits decreasing monotonicity from the $f_{k,l}$ family (since $2n_i > 2$ and $p_i > 0$). In addition, this function exhibits the following property, whose proof is given in the Appendix.

Property 1: The quotient $f(\rho)$ as defined in (7) boils down to a polynomial in z of degree Q

$$f(\rho) = \sum_{k=0}^{Q} F_k z^k \doteq F(z). \tag{8}$$

Moreover, the coefficient F_1 in (8) is always zero.

Now, one can construct SNR estimators by equating $F(\hat{z})$ to the sample-moments version of (7), i.e.,

$$F(\hat{z}) = \left(\prod_{i=1}^{U} \hat{M}_{2n_i}^{p_i}\right) \Big/ \hat{M}_2^Q \tag{9}$$

and then solving for $\hat{z}$ in (0,1). Monotonicity of $f(\rho)$ guarantees uniqueness of the solution. This procedure is in fact a polynomial root-finding problem, which may well be implemented by means of LUTs. Nevertheless, other less memory-costly choices are possible. On one hand, there exist standard direct algebraic solutions for the roots of polynomials of degree 4 or less ($2Q \leq 8$); on the other hand, iterative polynomial root-finding algorithms can be used, regardless of the order.

The popular $\mathrm{M_2M_4}$ estimator is a particular case of the family (7): it is a function of M_4/M_2^2, which by Property 1 reduces to $F(z) = (c_4 - 2)z^2 + 2$. The estimate is hence given by the positive root of $F(\hat{z}) = \hat{M}_4/\hat{M}_2^2$

$$\hat{z} = \sqrt{\frac{\hat{M}_4/\hat{M}_2^2 - 2}{c_4 - 2}}. \tag{10}$$

The quotient M_4/M_2^2 is the lowest-order ($2Q = 4$) member of this family. The next members, in growing statistical order, are M_6/M_2^3 ($2Q = 6$), M_4^2/M_2^4 and M_8/M_2^4 ($2Q = 8$), M_6M_4/M_2^5 and M_{10}/M_2^5 ($2Q = 10$), and so on.

III. A New Moments-Based Estimator

The desirable property of quotients of moments of the form (7) is that they can be reduced to a polynomial in z, so that estimator implementation becomes computationally simple. This applies as well to *any linear combination* of these quotients. This fact has been exploited in [14], where the linear combination $f(\rho) = M_6/M_2^3 - bM_4/M_2^2$ yields a sixth-order statistics-based estimate with improved behavior for multilevel constellations, as the weight b can be adjusted depending on the constellation. The good results in terms of variance attained by this estimator suggest that further improvements could be obtained by allowing higher-order quotients from the family (7) in the linear combination. Two questions arise at this point. First, how should the quotients featuring in the linear combination be selected? And second, once these quotients are somehow chosen, how should one select the weights?

Regarding the first question, there is no straightforward recipe for this "basis selection" problem. In practice, implementation complexity and finite precision effects would favor the selection of lower-order moments. On the other hand, including higher-order terms may be beneficial in terms of estimation bias and/or variance. Therefore, a tradeoff must be reached balancing these two conflicting goals. As for how to adjust the weights, it is sensible to optimize them in terms of statistical performance. This issue will be discussed in Section V.

Since the fourth- and sixth-order combinations have been already analyzed in [10] and [14] respectively, henceforth we focus on the weighted linear combination of *all* quotients up to order eight (EOS stands for *eighth-order statistics*)

$$f_{\mathrm{EOS}}(\rho) \doteq \beta\frac{M_4}{M_2^2} + \gamma\frac{M_6}{M_2^3} + \delta\frac{M_4^2}{M_2^4} + \epsilon\frac{M_8}{M_2^4} \tag{11}$$

for some weights $\boldsymbol{\alpha} = [\beta\ \gamma\ \delta\ \epsilon]^{\mathrm{T}} \in \mathbb{R}^4$. Note that $f_{\mathrm{EOS}}(\rho)$ includes as particular cases $\mathrm{M_2M_4}$ (by setting $\beta \neq 0$, $\gamma = \delta = \epsilon = 0$) and $\mathrm{M_6}$ ($\beta = -b$, $\gamma = 1$, and $\delta = \epsilon = 0$). For further reference we also define the vector $\boldsymbol{m} \doteq [M_2\ M_4\ M_6\ M_8]^{\mathrm{T}} \in \mathbb{R}^4$ containing the true moments appearing in (11), and the function $h_{\mathrm{EOS}}(\boldsymbol{m}) \doteq f_{\mathrm{EOS}}(\rho)$ considered from taking $\boldsymbol{m}$ as the independent variable in (11).

After applying the variable change $\rho = z/(1-z)$, one arrives at

$$F_{\mathrm{EOS}}(z) = F_4z^4 + F_3z^3 + F_2z^2 + F_0 \tag{12a}$$
$$= z^2(F_4z^2 + F_3z + F_2) + F_0 \tag{12b}$$

where the coefficients F_k are linear in the weights, and are given by

$$F_4 = \delta(c_4 - 2)^2 + \epsilon\,[72(c_4 - 1) - 16c_6 + c_8] \tag{13a}$$
$$F_3 = (\gamma + 16\epsilon)(12 - 9c_4 + c_6) \tag{13b}$$
$$F_2 = (\beta + 9\gamma + 4\delta + 72\epsilon)(c_4 - 2) \tag{13c}$$
$$F_0 = 2(\beta + 3\gamma + 2\delta + 12\epsilon). \tag{13d}$$

The estimator is found by solving for $\hat{z}$ in

$$F_{\mathrm{EOS}}(\hat{z}) = \beta\frac{\hat{M}_4}{\hat{M}_2^2} + \gamma\frac{\hat{M}_6}{\hat{M}_2^3} + \delta\frac{\hat{M}_4^2}{\hat{M}_2^4} + \epsilon\frac{\hat{M}_8}{\hat{M}_2^4} = h_{\mathrm{EOS}}(\hat{\boldsymbol{m}}) \tag{14}$$

where $\hat{\boldsymbol{m}} \doteq [\hat{M}_2\ \hat{M}_4\ \hat{M}_6\ \hat{M}_8]^{\mathrm{T}} \in \mathbb{R}^4$ is the vector of sample moments. This amounts to finding the roots of the quartic polynomial $F_{\mathrm{EOS}}(\hat{z}) - h_{\mathrm{EOS}}(\hat{\boldsymbol{m}})$, in which only the independent term $F_0 - h_{\mathrm{EOS}}(\hat{\boldsymbol{m}})$ depends on the observations. The roots of quartic polynomials can be algebraically found, although the procedure is somewhat intricate (see, e.g., [28, pp. 18–19]). Alternatively, one can resort to root-finding algorithms. For example, (12b) suggests the following iterative rule:

$$\hat{z}^{(0)} = 1, \quad \hat{z}^{(n+1)} = \sqrt{\frac{h_{\mathrm{EOS}}(\hat{\boldsymbol{m}}) - F_0}{F_4\left(\hat{z}^{(n)}\right)^2 + F_3\hat{z}^{(n)} + F_2}}. \tag{15}$$

Note that the polynomial $F_{\mathrm{EOS}}(z) - h_{\mathrm{EOS}}(\boldsymbol{m})$ should have a unique root in (0,1), i.e., $f_{\mathrm{EOS}}(\rho)$ should be monotonic. Given that the roots depend on the constellation moments c_p as well as on the weights $\boldsymbol{\alpha}$, one might ask whether conditions on $\boldsymbol{\alpha}$ can be given ensuring monotonicity of $f_{\mathrm{EOS}}(\rho)$; the answer to this question is not trivial. In the next sections, we show how $\boldsymbol{\alpha}$ can be optimized in terms of statistical performance, for a given constellation and SNR operation range. Let us advance that the optimal weights seem to yield monotonic functions $f_{\mathrm{EOS}}(\rho)$ in the majority of tested cases, with some exceptions to be discussed in Sections V and VI.

To close this section, let us remark that this EOS approach can be readily generalized to higher-order linear combinations of the form similar to (11). The presence of statistics of tenth or greater order will lead to polynomials $F(z)$ of degree 5 or greater. In such cases, direct computation of the root is not possible, but a root-finding rule of the form (15) always exists, since $F_1 = 0$ by Property 1.

IV. Statistical Analysis

Next we present a small-error analysis to obtain approximate expressions for the variance and bias of the EOS-based estimate introduced in Section III. Our goal is to expose the dependence of these performance measures with the weight vector $\boldsymbol{\alpha}$, in order to address weight optimization in Section V. We note that the analytical approach of this section can be readily generalized to estimators based on higher-order quotients of the form (7).

A. Variance

From (14), the estimator is given by $\hat{z} = F_{\mathrm{EOS}}^{-1}(h_{\mathrm{EOS}}(\hat{\boldsymbol{m}})) \doteq g(\hat{\boldsymbol{m}})$ and $\hat{\rho} = \hat{z}/(1-\hat{z}) = g(\hat{\boldsymbol{m}})/(1-g(\hat{\boldsymbol{m}})) \doteq t(\hat{\boldsymbol{m}})$, where g is implicitly given by $F_{\mathrm{EOS}}(g) - h_{\mathrm{EOS}}(\hat{\boldsymbol{m}}) = 0$. Following a standard procedure (see e.g. [25, Sec. 9.5]) we consider a first-order Taylor expansion of the estimator $\hat{\rho}$ about $\hat{\boldsymbol{m}} = \boldsymbol{m}$, which yields the approximation $\hat{\rho} \approx \rho + \boldsymbol{v}^{\mathrm{T}}(\hat{\boldsymbol{m}} - \boldsymbol{m})$, with $\boldsymbol{v} \doteq \nabla t|_{\hat{\boldsymbol{m}}=\boldsymbol{m}}$. Therefore

$$\mathrm{Var}\{\hat{\rho}\} \approx \boldsymbol{v}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{v} \tag{16}$$

with $\boldsymbol{C}$ the covariance matrix of $\hat{\boldsymbol{m}}$, whose elements are given by $[\boldsymbol{C}]_{ij} = (M_{2(i+j)} - M_{2i}M_{2j})/K$, $i,j \in \{1,2,3,4\}$. The computation of (16) is outlined in the Appendix; eventually, one arrives at

$$\mathrm{Var}\{\hat{\rho}\} \approx \frac{1}{K}\frac{A^{(\mathrm{Var})}(\rho)}{B^{(\mathrm{Var})}(\rho)} = \frac{1}{K}\frac{\sum_{n=0}^{10} A_n^{(\mathrm{Var})}\rho^n}{\sum_{n=2}^{6} B_n^{(\mathrm{Var})}\rho^n} \tag{17}$$

with $A^{(\mathrm{Var})}(\rho) \doteq \sum_{n=0}^{10} A_n^{(\mathrm{Var})}\rho^n$ and $B^{(\mathrm{Var})}(\rho) \doteq \sum_{n=2}^{6} B_n^{(\mathrm{Var})}\rho^n$. The coefficients of the former polynomials turn out to be quadratic in $\boldsymbol{\alpha}$

$$A_n^{(\mathrm{Var})} = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_n^{(\mathrm{Var})}\boldsymbol{\alpha} \Rightarrow A^{(\mathrm{Var})}(\rho) = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}^{(\mathrm{Var})}(\rho)\boldsymbol{\alpha} \tag{18}$$

$$B_n^{(\mathrm{Var})} = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{B}_n^{(\mathrm{Var})}\boldsymbol{\alpha} \Rightarrow B^{(\mathrm{Var})}(\rho) = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{B}^{(\mathrm{Var})}(\rho)\boldsymbol{\alpha} \tag{19}$$

where the 4×4 matrices $\boldsymbol{A}_n^{(\mathrm{Var})}$, $\boldsymbol{B}_n^{(\mathrm{Var})}$ are functions of the constellation moments only, and $\boldsymbol{A}^{(\mathrm{Var})}(\rho) = \sum_{n=0}^{10} \boldsymbol{A}_n^{(\mathrm{Var})}\rho^n$, $\boldsymbol{B}^{(\mathrm{Var})}(\rho) = \sum_{n=2}^{6} \boldsymbol{B}_n^{(\mathrm{Var})}\rho^n$. Note that (17) predicts that the variance is $\mathcal{O}(\rho^4)$ at high SNR, similarly to previous results for other estimators based on quotients of moments [10], [14].

The matrices in (18) and (19) exhibit interesting properties regardless of the constellation, which we summarize next for further reference. The proofs can be found in the Appendix.

Property 2: The 4×4 matrix $\boldsymbol{A}^{(\mathrm{Var})}(\rho)$ has rank not exceeding three.

Property 3: The 4×4 matrix $\boldsymbol{A}_{10}^{(\mathrm{Var})}$ has rank not exceeding three. In addition, $\mathrm{rank}(\boldsymbol{A}_{10}^{(\mathrm{Var})}) \le \min\{3, I-1\}$ (recall that I is the number of amplitude levels of the constellation).

Property 4: The 4×4 matrix $\boldsymbol{B}^{(\mathrm{Var})}(\rho)$ is of the form $\boldsymbol{B}^{(\mathrm{Var})}(\rho) = \boldsymbol{b}^{(\mathrm{Var})}(\rho)(\boldsymbol{b}^{(\mathrm{Var})}(\rho))^{\mathrm{T}}$ for some 4×1 nonzero vector $\boldsymbol{b}^{(\mathrm{Var})}(\rho)$, and therefore its rank is one.

Property 5: The 4×4 matrix $\boldsymbol{B}_6^{(\mathrm{Var})}$ is of the form $\boldsymbol{B}_6^{(\mathrm{Var})} = \boldsymbol{b}_6^{(\mathrm{Var})}(\boldsymbol{b}_6^{(\mathrm{Var})})^{\mathrm{T}}$ for some 4×1 nonzero vector $\boldsymbol{b}_6^{(\mathrm{Var})}$, and therefore its rank is one.

B. Bias

In order to obtain an approximation for the estimation bias, we use a a second-order Taylor expansion of $\hat{\rho}$ about $\hat{\boldsymbol{m}} = \boldsymbol{m}$: $\hat{\rho} \approx \rho + \boldsymbol{v}^{\mathrm{T}}(\hat{\boldsymbol{m}} - \boldsymbol{m}) + (1/2)(\hat{\boldsymbol{m}} - \boldsymbol{m})^{\mathrm{T}}\boldsymbol{H}(\hat{\boldsymbol{m}} - \boldsymbol{m})$, where $\boldsymbol{H}$ is the Hessian matrix of $t(\hat{\boldsymbol{m}})$ evaluated at $\hat{\boldsymbol{m}} = \boldsymbol{m}$. Straightforward algebra yields the approximation

$$\mathrm{Bias}\{\hat{\rho}\} \approx \frac{1}{2}\mathrm{Tr}\{\boldsymbol{H}\boldsymbol{C}\}. \tag{20}$$

After a few steps, as outlined in the Appendix, one finds

$$\mathrm{Bias}\{\hat{\rho}\} \approx \frac{1}{K}L(\rho) + J(\rho)\mathrm{Var}\{\hat{\rho}\} \tag{21a}$$

$$= \frac{1}{K}\frac{A^{(L)}(\rho)}{B^{(L)}(\rho)} + \frac{A^{(J)}(\rho)}{B^{(L)}(\rho)}\frac{1}{K}\frac{A^{(\mathrm{Var})}(\rho)}{B^{(\mathrm{Var})}(\rho)} \tag{21b}$$

$$= \frac{1}{K}\frac{\sum_{n=0}^{12} A_n^{(\mathrm{Bias})}\rho^n}{\sum_{n=3}^{9} B_n^{(\mathrm{Bias})}\rho^n} = \frac{1}{K}\frac{A^{(\mathrm{Bias})}(\rho)}{B^{(\mathrm{Bias})}(\rho)} \tag{21c}$$

where $L(\rho) \doteq A^{(L)}(\rho)/B^{(L)}(\rho)$, with $A^{(L)}(\rho) = \sum_{n=0}^{6} A_n^{(L)}\rho^n$, $B^{(L)}(\rho) = \sum_{n=1}^{4} B_n^{(L)}\rho^n$, and $J(\rho) \doteq A^{(J)}(\rho)/B^{(L)}(\rho)$, with $A^{(J)}(\rho) = \sum_{n=0}^{3} A_n^{(J)}\rho^n$ (note that $L(\rho)$ and $J(\rho)$ share the same denominator). Furthermore, the former polynomials and their coefficients are linear in $\boldsymbol{\alpha}$

$$A_n^{(L)} = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{a}_n^{(L)} \quad \Rightarrow \quad A^{(L)}(\rho) = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{a}^{(L)}(\rho) \tag{22}$$

$$B_n^{(L)} = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{b}_n^{(L)} \quad \Rightarrow \quad B^{(L)}(\rho) = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{b}^{(L)}(\rho) \tag{23}$$

$$A_n^{(J)} = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{a}_n^{(J)} \quad \Rightarrow \quad A^{(J)}(\rho) = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{a}^{(J)}(\rho) \tag{24}$$

where the 4×1 vectors $\boldsymbol{a}_n^{(L)}$, $\boldsymbol{b}_n^{(L)}$, $\boldsymbol{a}_n^{(J)}$ are functions of the constellation moments only, and $\boldsymbol{a}^{(L)}(\rho) = \sum_{n=0}^{6} \boldsymbol{a}_n^{(L)}\rho^n$, $\boldsymbol{b}^{(L)}(\rho) = \sum_{n=1}^{4} \boldsymbol{b}_n^{(L)}\rho^n$, $\boldsymbol{a}^{(J)}(\rho) = \sum_{n=0}^{3} \boldsymbol{a}_n^{(J)}\rho^n$. Therefore, $A^{(\mathrm{Bias})}$ and $B^{(\mathrm{Bias})}$ are *cubic* in $\boldsymbol{\alpha}$. Note that, at high SNR, $L(\rho)$ is $\mathcal{O}(\rho^2)$ whereas $J(\rho)$ is $\mathcal{O}(\rho^{-1})$, so (21) predicts that the bias is $\mathcal{O}(\rho^3)$ at high SNR.

C. MSE

Recalling that the minimum square error (MSE) of an estimator is given by $\mathrm{MSE}\{\hat{\rho}\} = \mathrm{Var}\{\hat{\rho}\} + \mathrm{Bias}^2\{\hat{\rho}\}$, and using (17) and (21), it is found that

$$\mathrm{MSE}\{\hat{\rho}\} \approx \frac{1}{K}\frac{\sum_{n=0}^{10} A_n^{(\mathrm{Var})}\rho^n}{\sum_{n=2}^{6} B_n^{(\mathrm{Var})}\rho^n} + \frac{1}{K^2}\left(\frac{\sum_{n=0}^{12} A_n^{(\mathrm{Bias})}\rho^n}{\sum_{n=3}^{9} B_n^{(\mathrm{Bias})}\rho^n}\right)^2. \tag{25}$$

Therefore the MSE is the sum of two terms, the first of which (the variance) is $\mathcal{O}(\rho^4)$ at high SNR and inversely proportional to K, and the second one (the squared bias) is $\mathcal{O}(\rho^6)$ at high SNR and inversely proportional to K^2. The dependence of the second term of (25) with $\boldsymbol{\alpha}$ is thus of *sixth order*.

V. Weight Optimization

We focus now on the selection of the weight vector $\boldsymbol{\alpha}$. Note that the four weights in $\boldsymbol{\alpha}$ only provide in fact *three* degrees of freedom, since the solutions to (14) are invariant to scalings in $\boldsymbol{\alpha}$. This is further emphasized by the fact that the approximations (17) and (21) for the variance and the bias are also invariant under this operation. In general, the number of degrees of freedom in this kind of estimators equals the number of quotients of moments in the linear combination minus one; for instance, M_6 has one degree of freedom, whereas $\mathrm{M}_2\mathrm{M}_4$ has none. We present two weight selection procedures, according to two different optimization criteria. Ideally, one would like to obtain an unbiased estimator with variance close to the theoretical limit dictated by the Cramér-Rao Bound (CRB). This cannot be achieved with the limited degrees of freedom available, and therefore one should settle for a less ambitious goal. The question that arises then is how to trade off bias and variance, and where (i.e., in which SNR operating region).

For example, in [14] the only free parameter in M_6 was optimized to yield minimum variance in the high SNR region. This approach can be generalized to the case in which several degrees of freedom are available, taking the bias into account as well; we refer to this criterion as "C1". Yet other approaches are possible. For instance, systems using adaptive coding and modulation commonly use each particular constellation within a limited SNR range (see, e.g., [29]). It seems then reasonable to optimize the performance for some nominal SNR within that range, with the hope that the estimator will still perform well in a neighborhood of this nominal value. We will refer to this criterion as "C2". Note that either C1 or C2 could reasonably be applied to any estimator derived from a linear combination of the form of (11).

A. Criterion C1: Weight Optimization for High SNR

Asymptotically as $\rho \to \infty$, all polynomial divisions $A^{(\cdot)}(\rho)/B^{(\cdot)}(\rho)$ in (17) and (21) can be accurately approximated by the first few terms of the quotient polynomial

$$\begin{aligned}\mathrm{Var}\{\hat{\rho}\} &\approx \frac{1}{K}\frac{A^{(\mathrm{Var})}(\rho)}{B^{(\mathrm{Var})}(\rho)} \doteq \frac{1}{K}C(\rho)\\ &\approx \frac{1}{K}\left(C_4\rho^4 + C_3\rho^3 + C_2\rho^2 + C_1\rho + C_0\right)\end{aligned} \tag{26}$$

$$L(\rho) \approx L_2\rho^2 + L_1\rho + L_0 \tag{27}$$

$$J(\rho) \approx J_{-1}\rho^{-1} + J_{-2}\rho^{-2} + J_{-3}\rho^{-3}. \tag{28}$$

It makes sense now to use the available degrees of freedom to minimize (or, if possible, cancel) the magnitude of the highest-order coefficients of the variance and the terms of the bias.[4] As for the bias, see (21a), it seems convenient to separately minimize $L(\rho)$, $J(\rho)$ and the variance. In this regard, it must be noted that $J_{-1} = A_3^{(J)}/B_4^{(L)} = 2$ independently of $\boldsymbol{\alpha}$ and the constellation. This suggests that the minimization efforts should focus on $L(\rho)$ and the variance. Regarding $L(\rho)$, its highest-order coefficient is given by

$$L_2 = \frac{A_6^{(L)}}{B_4^{(L)}} = \frac{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{a}_6^{(L)}}{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{b}_4^{(L)}}. \tag{29}$$

Observe that unless $\boldsymbol{a}_6^{(L)}$ and $\boldsymbol{b}_4^{(L)}$ are colinear (which is not the case with typical constellations), (29) can be made zero. Thus, we propose the general principle for weight optimization in the high SNR region.

Criterion C1: Spend one degree of freedom in $\boldsymbol{\alpha}$ to achieve $L_2 = 0$. Spend the remaining degrees of freedom in order to minimize C_4. If C_4 can be made zero and there are degrees of freedom to spare, then proceed to minimize C_3, and so on.

With this approach, $L(\rho)$ becomes $\mathcal{O}(\rho)$. Note that $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{a}_6^{(L)} = 0$ implies that $\boldsymbol{\alpha} = \boldsymbol{F}\bar{\boldsymbol{\alpha}}$ for some 3×1 vector $\bar{\boldsymbol{\alpha}}$, where $\boldsymbol{F}$ is a 4×3 matrix satisfying $\boldsymbol{F}^{\mathrm{T}}\boldsymbol{F} = \boldsymbol{I}$ and $\boldsymbol{F}^{\mathrm{T}}\boldsymbol{a}_6^{(L)} = \boldsymbol{0}$. The columns of $\boldsymbol{F}$ can be selected as the three eigenvectors associated to the three nonzero eigenvalues of the projection matrix

$$\boldsymbol{W} \doteq \boldsymbol{I} - \frac{\boldsymbol{a}_6^{(L)}\left(\boldsymbol{a}_6^{(L)}\right)^{\mathrm{T}}}{\left\|\boldsymbol{a}_6^{(L)}\right\|^2}. \tag{30}$$

Now, the highest-order coefficient of the variance is given by [cf. (17)–(19) and Property 5]

$$C_4 = \frac{A_{10}^{(\mathrm{Var})}}{B_6^{(\mathrm{Var})}} = \frac{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_{10}^{(\mathrm{Var})}\boldsymbol{\alpha}}{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{b}_6^{(\mathrm{Var})}\left(\boldsymbol{b}_6^{(\mathrm{Var})}\right)^{\mathrm{T}}\boldsymbol{\alpha}}. \tag{31}$$

Thus, for $\boldsymbol{\alpha} = \boldsymbol{F}\bar{\boldsymbol{\alpha}}$, C_4 becomes

$$C_4 = \frac{\bar{\boldsymbol{\alpha}}^{\mathrm{T}}\boldsymbol{P}\bar{\boldsymbol{\alpha}}}{\bar{\boldsymbol{\alpha}}^{\mathrm{T}}\boldsymbol{q}\boldsymbol{q}^{\mathrm{T}}\bar{\boldsymbol{\alpha}}} \tag{32}$$

where $\boldsymbol{P} \doteq \boldsymbol{F}^{\mathrm{T}}\boldsymbol{A}_{10}^{(\mathrm{Var})}\boldsymbol{F}$ and $\boldsymbol{q} \doteq \boldsymbol{F}^{\mathrm{T}}\boldsymbol{b}_6^{(\mathrm{Var})}$. Note that (32) is a generalized Rayleigh quotient with a rank-one denominator matrix. It can be checked that the solution $\bar{\boldsymbol{\alpha}}_\star$ to the minimization of (32) (subject to $\bar{\boldsymbol{\alpha}}^{\mathrm{T}}\boldsymbol{q} \neq \boldsymbol{0}$) is as follows.

- If $\boldsymbol{P}$ is invertible, then $\bar{\boldsymbol{\alpha}}_\star = \boldsymbol{P}^{-1}\boldsymbol{q}$. The minimized value of C_4 is strictly positive.

[4]Alternatively, one could attempt to minimize the highest-order coefficients of the high-SNR approximation of the MSE. We do not follow this approach, however, in view of the sixth-order dependence with the weights of some of the terms in (25).

TABLE I
OPTIMAL WEIGHTS ($\epsilon = 1$) UNDER CRITERION C1. APSK CONSTELLATIONS ARE AS DEFINED IN [29] FOR THE SPECIFIED CODE RATES

Constellation	I	β	γ	δ
16-APSK (2/3)	2	5.9396	−2.8400	−1.4325
16-APSK (3/4)	2	6.0768	−2.8769	−1.4572
16-APSK (4/5)	2	6.1331	−2.8918	−1.4676
16-APSK (5/6)	2	6.1637	−2.8999	−1.4734
16-APSK (8/9)	2	6.2306	−2.9173	−1.4860
16-APSK (9/10)	2	6.2522	−2.9229	−1.4901
32-APSK (3/4)	3	12.3187	−3.8576	−2.7445
32-APSK (4/5)	3	12.0919	−3.8541	−2.7024
32-APSK (5/6)	3	11.9647	−3.8527	−2.6762
32-APSK (8/9)	3	11.6972	−3.8501	−2.6102
32-APSK (9/10)	3	11.6684	−3.8498	−2.6027
16-QAM	3	9.9411	−5.28	−0.2807
32-QAM	5	10.2400	−3.8552	−2.1227
64-QAM	9	10.7991	−4.3509	−1.8525
128-QAM	16	10.7081	−4.1297	−2.0170
256-QAM	32	10.4846	−4.4251	−1.6505

- If $\boldsymbol{P}$ is singular and $\mathrm{rank}([\boldsymbol{P}\ \boldsymbol{q}]) = \mathrm{rank}(\boldsymbol{P})$, then there exist infinitely many solutions, but all of them yield the same minimized value of C_4, which again is strictly positive. One of these solutions is $\bar{\boldsymbol{\alpha}}_\star = \boldsymbol{P}^{\#}\boldsymbol{q}$, where $\boldsymbol{P}^{\#}$ denotes the pseudoinverse of $\boldsymbol{P}$.
- If $\boldsymbol{P}$ is singular and $\mathrm{rank}([\boldsymbol{P}\ \boldsymbol{q}]) > \mathrm{rank}(\boldsymbol{P})$, then C_4 can be made zero. The solutions can be given in terms of the singular value decomposition (SVD) of $\boldsymbol{P}$

$$\boldsymbol{P} = [\boldsymbol{U} \quad \boldsymbol{V}] \begin{bmatrix} \boldsymbol{\Sigma} & \\ & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{U}^{\mathrm{T}} \\ \boldsymbol{V}^{\mathrm{T}} \end{bmatrix} \tag{33}$$

where $\boldsymbol{\Sigma}$ is positive definite, and $\boldsymbol{U}$, $\boldsymbol{V}$ have orthonormal columns. Then all the vectors $\bar{\boldsymbol{\alpha}}_\star$ such that $\boldsymbol{P}\bar{\boldsymbol{\alpha}}_\star = \boldsymbol{0}$ and $\bar{\boldsymbol{\alpha}}_\star^{\mathrm{T}}\boldsymbol{q} \neq \boldsymbol{0}$ are given (up to a scaling) by

$$\bar{\boldsymbol{\alpha}}_\star = \boldsymbol{V}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{q} + \boldsymbol{V}\boldsymbol{V}^{\mathrm{T}}\left(\boldsymbol{I} - \frac{\boldsymbol{q}\boldsymbol{q}^{\mathrm{T}}}{\|\boldsymbol{V}^{\mathrm{T}}\boldsymbol{q}\|^2}\right)\boldsymbol{V}\boldsymbol{s} \tag{34}$$

where $\boldsymbol{s}$ is arbitrary.

Let $r \doteq \mathrm{rank}(\boldsymbol{A}_{10}^{(\mathrm{Var})})$, and note that $\mathrm{rank}(\boldsymbol{P}) \leq r$. Although in general $\mathrm{rank}(\boldsymbol{P})$ could be strictly smaller than r, for all constellations tested it holds that $\mathrm{rank}(\boldsymbol{P}) = r$. Also, recall that $r \leq \min\{3, I-1\}$ by Property 3. These facts imply that, depending on the number I of amplitude levels in the constellation, one has the following.

- If $I = 1$ (CM constellations), then C_4 is automatically zero.
- If $I = 2$, then C_4 can be made zero spending just one degree of freedom.
- If $I = 3$, then C_4 can be made zero spending two degrees of freedom.
- If $I > 3$, then C_4 cannot be made zero.

Thus, estimation performance is expected to degrade as the number of levels in the constellation increases. More details are given on each specific case (in terms of I) in Section VI. For reference, Table I lists the optimal weights $\boldsymbol{\alpha}_\star = \boldsymbol{F}\bar{\boldsymbol{\alpha}}_\star$ under Criterion C1 for a number of practical multilevel constellations.[5]

Criterion C1 provides valid estimators for almost all practical constellations, in the sense that the resulting function $f_{\mathrm{EOS}}(\rho)$ in (11) is monotonic; the performance obtained is good in the medium to high SNR range. The only exception we have found is the 16-quadrature amplitude modulation (QAM) constellation, for which the weights provided by Criterion C1 cause the MSE to exhibit a sharp peak at $\rho \approx 19$ dB (the denominators $\boldsymbol{\alpha}_\star^{\mathrm{T}}\boldsymbol{b}^{(\mathrm{Var})}(\rho)$ and $\boldsymbol{\alpha}_\star^{\mathrm{T}}\boldsymbol{b}^{(L)}(\rho)$ become zero for this SNR value). Thus, Criterion C1 is not well suited to 16-QAM.

B. Criterion C2: Weight Optimization for a Nominal SNR Value

Considering a target SNR ρ_0 and using Property 4, one can write (21) and (17) as

$$\begin{aligned} \mathrm{Bias}\{\hat{\rho}\}|_{\rho=\rho_0} &\approx \frac{1}{K}L(\rho_0) + J(\rho_0)\,\mathrm{Var}\{\hat{\rho}\}|_{\rho=\rho_0} & (35)\\ &= \frac{1}{K}\frac{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{a}^{(L)}(\rho_0)}{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{b}^{(L)}(\rho_0)} \\ &\quad + \frac{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{a}^{(J)}(\rho_0)}{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{b}^{(L)}(\rho_0)}\,\mathrm{Var}\{\hat{\rho}\}|_{\rho=\rho_0} & (36) \end{aligned}$$

$$\mathrm{Var}\{\hat{\rho}\}|_{\rho=\rho_0} \approx \frac{1}{K}\frac{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}^{(\mathrm{Var})}(\rho_0)\boldsymbol{\alpha}}{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{b}^{(\mathrm{Var})}(\rho_0)\left[\boldsymbol{b}^{(\mathrm{Var})}(\rho_0)\right]^{\mathrm{T}}\boldsymbol{\alpha}}. \tag{37}$$

Therefore, it is natural to consider the following way of selecting the weights.

Criterion C2: Given a target SNR value ρ_0, minimize the approximate variance (37) in terms of $\boldsymbol{\alpha}$, under the constraint that $L(\rho_0)$ in (35) be zero.

This criterion amounts to minimizing a generalized Rayleigh quotient with a rank-one denominator matrix, subject to a linear constraint. This is analogous to the case encountered for Criterion C1. The only difference is that now the objective function (37) cannot be made zero, as it corresponds with the estimator variance (and not just one coefficient in its Taylor series expansion). Therefore, the solution is given (up to a scaling) by

$$\boldsymbol{\alpha}_\star(\rho_0) = \boldsymbol{F}\left(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{A}^{(\mathrm{Var})}(\rho_0)\boldsymbol{F}\right)^{-1}\boldsymbol{F}^{\mathrm{T}}\boldsymbol{b}^{(\mathrm{Var})}(\rho_0) \tag{38}$$

where the columns of the 4×3 matrix $\boldsymbol{F}$ are now given by the three eigenvectors associated to the three nonzero eigenvalues of

$$\boldsymbol{W} = \boldsymbol{I} - \frac{\boldsymbol{a}^{(L)}(\rho_0)\left[\boldsymbol{a}^{(L)}(\rho_0)\right]^{\mathrm{T}}}{\left\|\boldsymbol{a}^{(L)}(\rho_0)\right\|^2}. \tag{39}$$

This SNR-dependent solution $\boldsymbol{\alpha}_\star(\rho_0)$ proves in most cases competitive within a few decibels about ρ_0. Nevertheless, one must evaluate its performance in the entire SNR range for each particular constellation and target SNR, as $\boldsymbol{\alpha}_\star(\rho_0)$ could yield an invalid estimator in some SNR interval (i.e., $f_{\mathrm{EOS}}(\rho)$ in (11)

[5]The *Mathematica* and MATLAB code used to derive these optimal weights is available online at [30].

could turn out to be non-monotonic). Of all constellations tested, only the 16-APSK (Amplitude and Phase Shift Keying) family suffers from this problem, but only within some limited SNR intervals, as shown in the next section.

Note that in (35) and (36) the factor $J(\rho_0)$ has the same form as $L(\rho_0)$, and therefore, in contrast to Criterion C1, it could be made zero using one degree of freedom in $\boldsymbol{\alpha}$. This suggests an alternative criterion, in which one of the degrees of freedom used for minimizing the variance under Criterion C2 would be used instead for canceling $J(\rho_0)$ (and therefore completely canceling the bias at ρ_0). However, the solutions provided by this alternative criterion turn out to be problematic in practice ($f_{\mathrm{EOS}}(\rho)$ is not monotonic for most constellations; sharp peaks appear in bias and variance at certain SNRs), and hence it will not be further considered here.

VI. Performance Results

The performance of the proposed estimators depends strongly on the number of levels of the constellation. Next we discuss the results achieved with criteria C1 and C2 with respect to existing estimates of the same kind: $\mathrm{M_2M_4}$ and $\mathrm{M_6}$. We show results for the dependence of the bias with the SNR, and the dependence of the SNR-normalized MSE with the SNR and the samples size K. The SNR-normalized MSE is defined as $\mathrm{NMSE}\{\hat{\rho}\} \doteq \mathrm{MSE}\{\hat{\rho}\}/\rho^2$. The NDA-EVB CRB (numerically evaluated as in [16]) is provided as benchmark. Analytical results are completed with empirical results obtained through simulations, in which each point was averaged over 10 000 realizations.

A. CM Constellations

CM constellations (for which $c_p = 1$, for all p) constitute a special case, as it turns out that $A_6^{(L)} = 0$ whereas $A_5^{(L)} \neq 0$ (with $B_4^{(L)} \neq 0$), and that $A_{10}^{(\mathrm{Var})} = A_9^{(\mathrm{Var})} = 0$ whereas $A_8^{(\mathrm{Var})} \neq 0$ (with $B_6^{(\mathrm{Var})} \neq 0$). Therefore, the bias and variance are respectively linear and quadratic in ρ for general $\boldsymbol{\alpha}$. Interestingly, the highest-order coefficients of $L(\rho)$ and of the variance turn out to be independent of $\boldsymbol{\alpha}$: $A_5^{(L)}/B_4^{(L)} = 1$ and $A_8^{(\mathrm{Var})}/B_6^{(\mathrm{Var})} = 2$. It makes sense then to use a modified version of Criterion C1, focusing for example on the lowest-order coefficient of the variance $A_0^{(\mathrm{Var})}/B_2^{(\mathrm{Var})}$, which dominates in the region of *low* SNR. This term is minimized for $\boldsymbol{\alpha}_\star = [\beta\, 0\, \delta\, 0]^{\mathrm{T}}$, where β and δ cannot be both zero; the attained variance is $\mathrm{Var}\{\hat{\rho}\} \approx (2\rho^4 + 8\rho^3 + 10\rho^2 + 6\rho + 1)/(K\rho^2)$, independently of β and δ, and coincides with that of the $\mathrm{M_2M_4}$ estimator for CM constellations [10]. The predicted MSE turns out to remain quite close to the CRB throughout the entire SNR range.

B. Two-Level Constellations

When $I = 2$, application of Criterion C1 with the available degrees of freedom results in $L_2 = C_4 = C_3 = 0$. In this case, it is possible to obtain closed-form expressions for the optimal weights in terms of the probabilities $p = P_1$, $1 - p = P_2$, and the ring ratio $w = R_2/R_1$

$$\begin{aligned}
\beta =& \frac{(1+w^2)}{\left[p^2 - (1-p)^2 w^4\right]^2} \\
&\cdot \left[3w^6 + w^4 + 3p^2(1+w^2)(1+w^4) \right.\\
&\left. -2p(3w^6 + 2w^4 + w^2)\right] \cdot \epsilon, \\
\gamma =& -\frac{2(1+w^2)}{p + (1-p)w^2} \cdot \epsilon, \\
\delta =& -\frac{p + (1-p)w^4}{\left[p - (1-p)w^2\right]^2} \cdot \epsilon, \quad \epsilon \neq 0.
\end{aligned}$$

As a result, Criterion C1 achieves $\mathcal{O}(\rho^2)$ in the variance and the MSE, which is the lowest possible order as dictated by the CRB [16]. In contrast, the variances of $\mathrm{M_6}$ and $\mathrm{M_2M_4}$ are respectively $\mathcal{O}(\rho^3)$ and $\mathcal{O}(\rho^4)$, and their respective MSEs are $\mathcal{O}(\rho^4)$ and $\mathcal{O}(\rho^6)$.

Fig. 1(a) and (b) show the bias and the NMSE versus the SNR (for $K = 1000$ samples) for one of the 16-APSK constellations specified in [29] ($P_1 = 1/4$, $P_2 = 3/4$, $R_2/R_1 = 3.15$). The bias under Criterion C1 remains small, and the improvement in NMSE at high SNR is evident. Still, the gap to the CRB is not negligible. Note that for $\rho < 22$ dB, $\mathrm{M_6}$ achieves a slightly lower NMSE than Criterion C1, but its performance quickly degrades beyond this point.

Results for Criterion C2 with $\rho_0 = 20$ dB are also displayed in Fig. 1(b), showing that this is in fact a good strategy in the vicinity of ρ_0 (it outperforms the rest of estimators between 15 and 22 dB), but also for $\rho < \rho_0$, where its NMSE is never larger than that provided by Criterion C1 (this seems to hold regardless of the value of ρ_0). From a designer's perspective, setting the target SNR ρ_0 near the upper limit of the SNR operation range seems an appropriate choice.

For reference, Fig. 1(b) also shows the "genie-aided" NMSE curve, obtained by choosing the optimal $\boldsymbol{\alpha}_\star(\rho)$ under Criterion C2 for each ρ. This gives a lower bound on the NMSE for the proposed eighth-order estimator. Note that the genie-aided curve approaches the CRB as the SNR increases. Hence, within this region, the estimator designed under Criterion C2 is near optimal at the target SNR. In the low SNR region (below 5 dB), $\mathrm{M_2M_4}$ attains the genie-aided curve; however, neither $\mathrm{M_2M_4}$, nor $\mathrm{M_6}$, nor Criterion C1 can reach genie-aided performance above 5 dB. In general, good agreement is observed between theory and simulations.

As mentioned in Section V-B, Criterion C2 may yield invalid estimators for certain values of ρ_0. For the 16-APSK constellation of Fig. 1(b), this happens for ρ_0 between 5 and 15 dB. In this case one could fall back on some suboptimal solution, e.g., use the valid C2 estimator for $\rho_0 = 15$ dB, which offers good performance (not shown in Fig. 1) in this SNR interval.

Fig. 1(c) shows the variation of the NMSE with the number of observed samples K, for $\mathrm{SNR} = 20$ dB. It is seen that, for all estimators, the empirical results agree with the predicted theoretical values as long as the corresponding NMSE is sufficiently small (which is an indicator of the accuracy of the Taylor series approximation used to obtain the theoretical expressions).

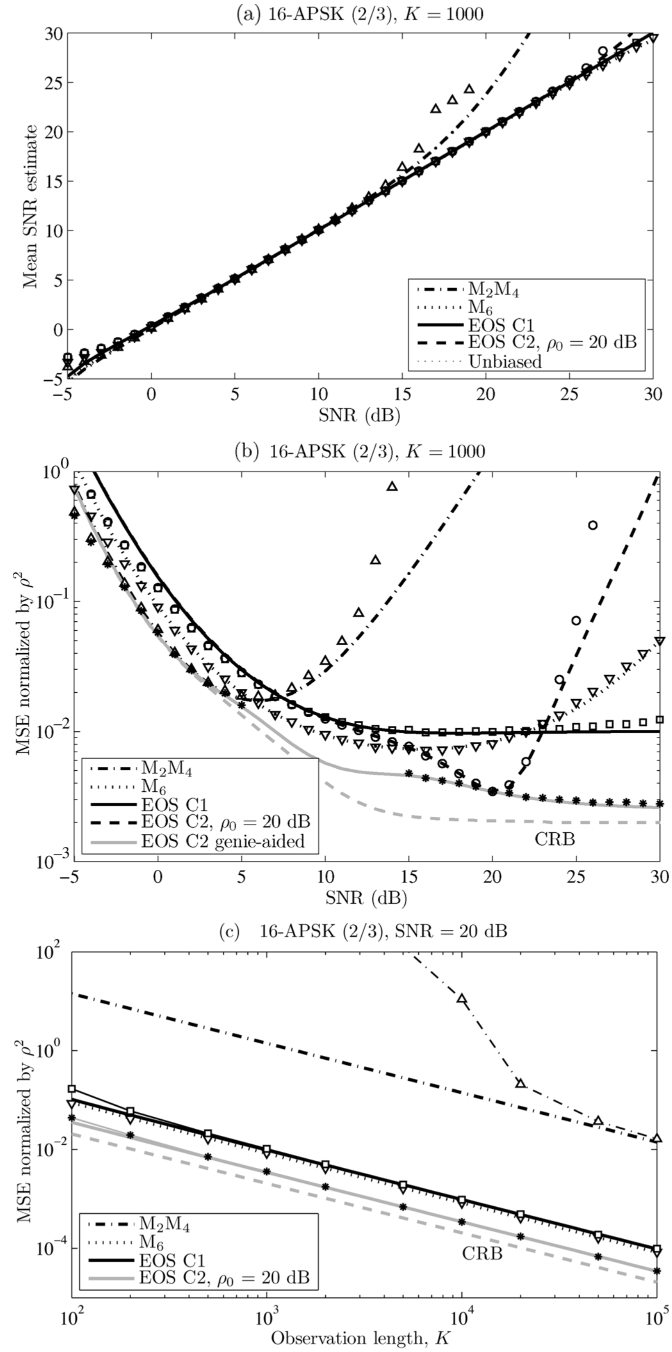

Fig. 1. Theoretical (lines) and empirical (markers) performance of the proposed estimator in terms of (a) bias versus SNR, (b) NMSE versus SNR, and (c) NMSE versus K for a 16-APSK with ring ratio $R_2/R_1 = 3.15$. Markers are as follows: $\triangle$ for $\mathrm{M}_2\mathrm{M}_4$, $\triangledown$ for M_6, $\square$ for C1, $\circ$ for C2 with $\rho_0 = 20$ dB, and $*$ for C2 genie-aided.

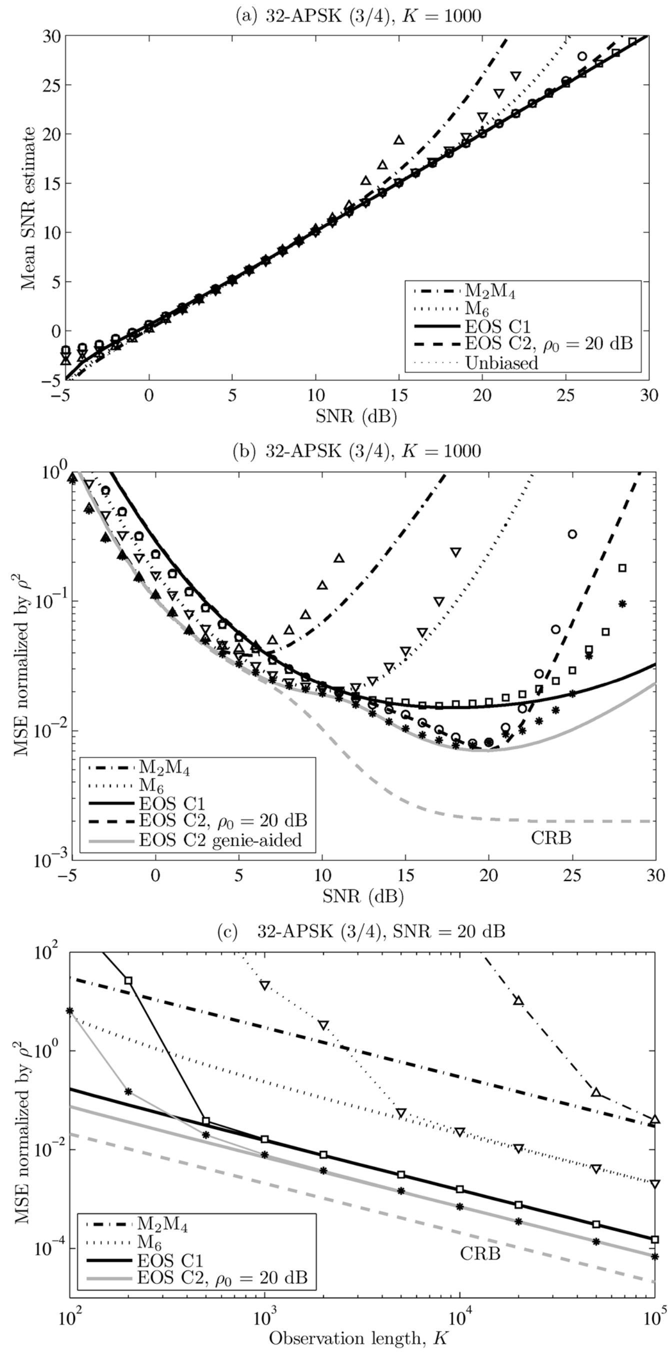

Fig. 2. Theoretical (lines) and empirical (markers) performance of the proposed estimator in terms of (a) bias versus SNR, (b) NMSE versus SNR, and (c) NMSE versus K for a 32-APSK with ring ratios $R_2/R_1 = 2.84$, $R_3/R_1 = 5.27$. Markers are as follows: $\triangle$ for $\mathrm{M}_2\mathrm{M}_4$, $\triangledown$ for M_6, $\square$ for C1, $\circ$ for C2 with $\rho_0 = 20$ dB, and $*$ for C2 genie-aided.

Note that the new EOS estimators outperform the $\mathrm{M}_2\mathrm{M}_4$ and M_6 schemes in the sense that they can achieve the same NMSE with significantly fewer samples. This holds true also for other SNR values.

C. Three-Level Constellations

For $I = 3$, application of Criterion C1 with the available degrees of freedom results in $L_2 = C_4 = 0$. The resulting variance and MSE are respectively $\mathcal{O}(\rho^3)$ and $\mathcal{O}(\rho^4)$, a remarkable improvement over $\mathrm{M}_2\mathrm{M}_4$ and M_6, which are both $\mathcal{O}(\rho^4)$ (variance) and $\mathcal{O}(\rho^6)$ (MSE). This can be seen in Fig. 2(a) and (b), which display the bias, and the NMSE versus the SNR (for $K = 1000$ samples) for one of the 32-APSK constellation specified in [29] ($P_1 = 1/4$, $P_2 = 3/8$, $P_3 = 1/2$, $R_2/R_1 = 2.84$, $R_3/R_1 = 5.27$). Good agreement is observed between theory and simulations. M_6 and Criterion C2 with $\rho_0 = 20$ dB remain close to the genie-aided case between 5 and 10 dB, and between 12 and 21 dB, respectively. Note that the genie-aided curve does

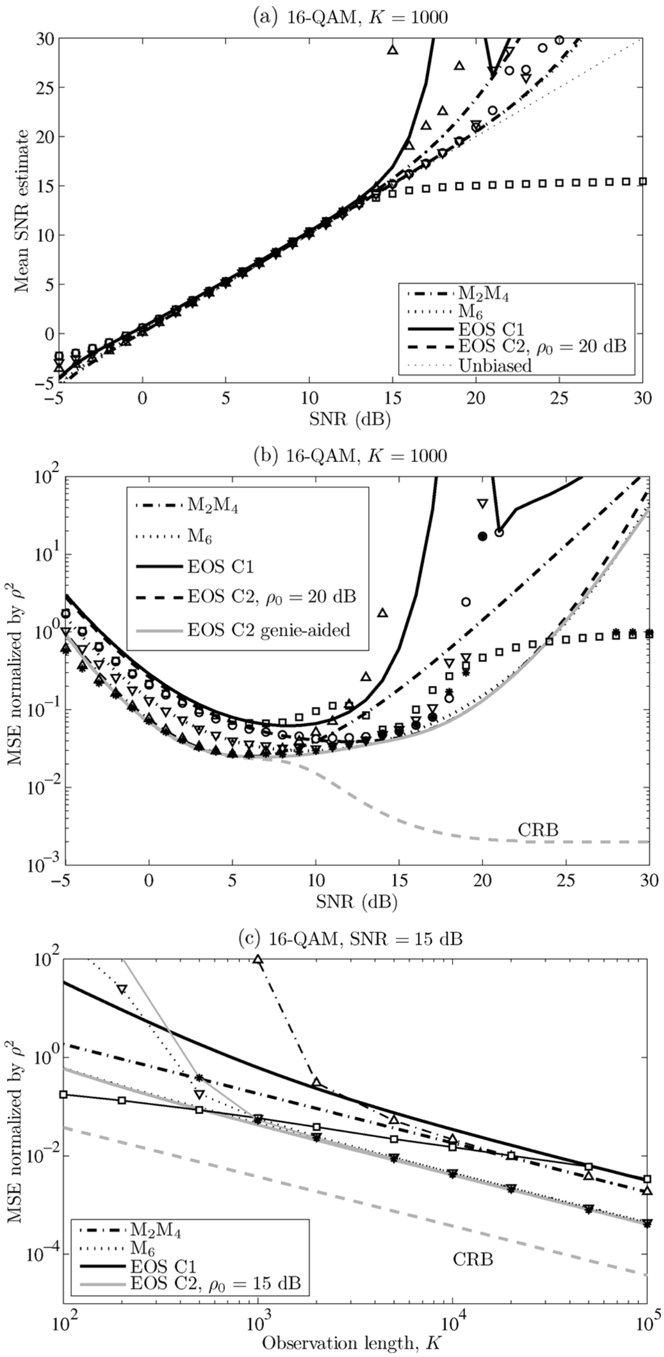

Fig. 3. Theoretical (lines) and empirical (markers) performance of the proposed estimator in terms of (a) bias versus SNR, (b) NMSE versus SNR, and (c) NMSE versus K for 16-QAM. Markers are as follows: $\triangle$ for $\mathrm{M_2M_4}$, ∇ for $\mathrm{M_6}$, $\square$ for C1, $\circ$ for C2 with $\rho_0 = 20$ dB, and $*$ for C2 genie-aided.

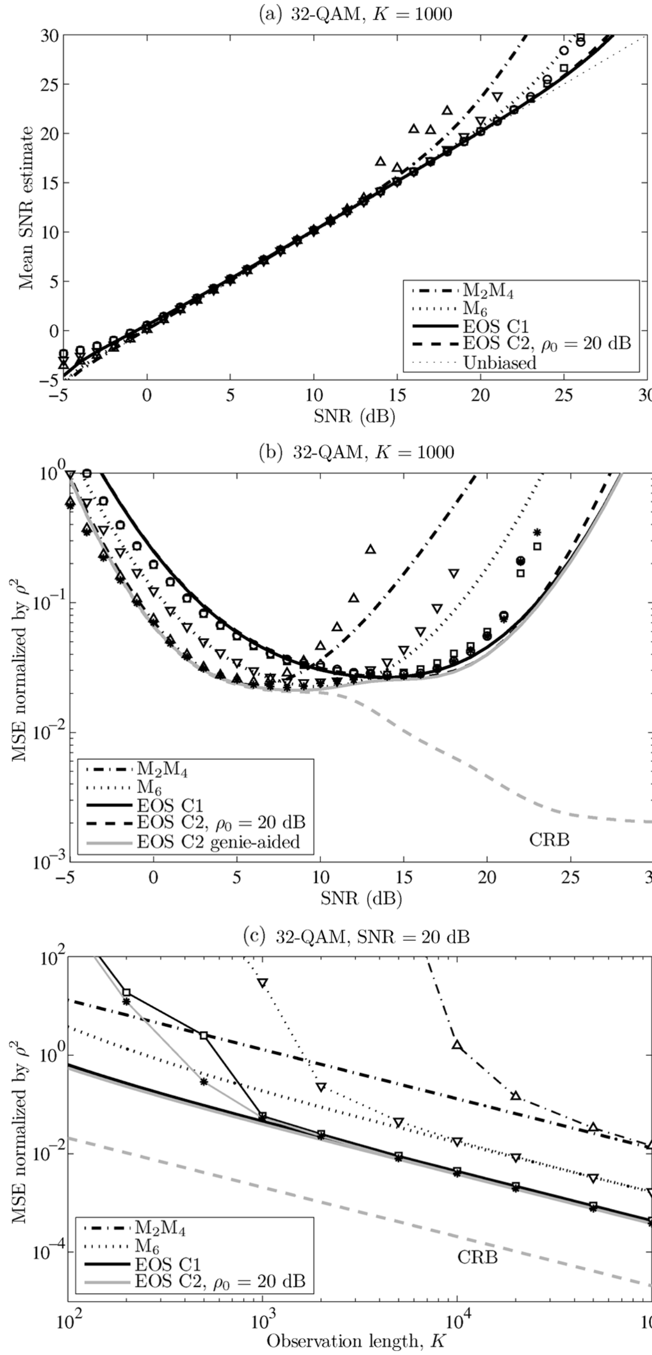

Fig. 4. Theoretical (lines) and empirical (markers) performance of the proposed estimator in terms of (a) bias versus SNR, (b) NMSE versus SNR, and (c) NMSE versus K for 32-QAM. Markers are as follows: $\triangle$ for $\mathrm{M_2M_4}$, ∇ for $\mathrm{M_6}$, $\square$ for C1, $\circ$ for C2 with $\rho_0 = 20$ dB, and $*$ for C2 genie-aided.

not approach the CRB in this case. Fig. 3(c) displays the dependence of the NMSE with K (for $\mathrm{SNR} = 20$ dB). As observed for 16-APSK, the theoretical analysis proves useful for sufficiently small NMSE, and the EOS estimators outperform $\mathrm{M_2M_4}$ and $\mathrm{M_6}$ in terms of K.

As mentioned in Section V-A, Criterion C1 fails to provide a good estimator for the 16-QAM constellation, see Fig. 3. The estimator designed under Criterion C1 is severely biased above 14 dB, and below that SNR its NMSE is outperformed by those of $\mathrm{M_2M_4}$ and $\mathrm{M_6}$ (observe also the peak in the theoretical NMSE at about 19 dB). From Fig. 3(b) and (c) it is seen that $\mathrm{M_6}$ almost achieves genie-aided performance above 10 dB, which suggests that the EOS approach is not advantageous for 16-QAM. Comparison with Fig. 2 shows that the best NMSEs obtained at intermediate-to-high SNR for 16-QAM are substantially higher than those obtained for 32-APSK, *for all moments-based methods*. Thus, the arrangement of signal points for 16-QAM seems to be particularly challenging to this class of estimators.

D. Constellations With More Than Three Levels

For $I \geq 4$, application of Criterion C1 with the available degrees of freedom results in $L_2 = 0$ and $C_4 > 0$. The resulting variance and MSE are therefore $\mathcal{O}(\rho^4)$ and $\mathcal{O}(\rho^6)$, respectively. Nevertheless, this approach is still useful, as shown in Fig. 4, which shows the results obtained with 32-QAM ($I = 5$). In terms of NMSE, (Fig. 4(b), $K = 1000$ samples), Criterion C1 clearly improves over $\mathrm{M_2M_4}$ and $\mathrm{M_6}$ in the high SNR region. $\mathrm{M_2M_4}$ achieves close-to-genie-aided performance below 6 dB, whereas $\mathrm{M_6}$ takes over between 7 and 13 dB. For higher SNR, criteria C1 and C2 ($\rho_0 = 20$ dB), which show almost identical performance, are the most competitive. The results in terms of K [see Fig. 4(c)] confirm the same trend observed for 16-APSK and 32-APSK.

The behavior for other QAM constellations with more than three levels is similar to that of Fig. 4. We must note that the performance improvement obtained with eighth-order estimators at high SNR is observed to be greater for cross-QAM constellations (32-, 128-QAM) than for square QAMs (64-, 256-QAM).

VII. Conclusion

Due to its simplicity, moments-based SNR estimation is an attractive choice, which until recently proved competitive only for constant modulus constellations. Recent efforts tried to extend its application to higher-order modulations under different approaches. One of them is to include higher-order statistics in the computation of the estimates. We developed an SNR estimator based on eighth-order statistics, which can be efficiently implemented without LUTs. The estimator is built upon a linear combination of quotients of moments, whose weights can be tuned according to the constellation and the SNR operation range. Two weight optimization criteria are proposed, which yield good estimators with improved performance over existing methods, particularly for two- and three-level constellations. Possible extensions of the proposed estimator to higher orders (e.g., extending the linear combination with tenth-order statistics) would presumably bring further performance improvements. The analysis and optimization of such extensions can be carried out using the approach that was applied here to the eighth-order estimator.

Appendix

Here we provide the proofs of the results given in this paper. Supporting *Mathematica* code for some of the algebraic derivations is available online at [30].

A. Proof of Property 1

Recalling (4), it is possible to write the numerator and denominator of (7) as

$$\prod_{i=1}^{U} M_{2n_i}^{p_i} = N^Q D(\rho) \tag{40}$$

$$M_2^Q = N^Q (1+\rho)^Q \tag{41}$$

with $D(\rho)$ a polynomial of degree Q. As expected, the ratio of (40) to (41) does not depend on N. Now, applying the change of variable $\rho = z/(1-z)$, one has $\rho^m/(1+\rho)^Q = z^m(1-z)^{Q-m}$, for $m = 0, 1, \ldots, Q$, which shows that

$$f(\rho) = \sum_{m=0}^{Q} D_m z^m (1-z)^{Q-m} = F(z) \tag{42}$$

is a polynomial of degree Q in z. We note that this property holds also for any linear combination of quotients of the form of (7). We will prove now that F_1, the coefficient in z of $F(z) = f(\rho)$, is always zero. Since $f(\rho)$ is a product of terms of the form M_{2n}/M_2^n, it suffices to show that any such term is a polynomial in z with zero first-order coefficient. To see this, note from (4) and $\rho = z/(1-z)$ that $M_{2n}/M_2^n = \sum_{m=0}^{n} G_m^{(n)} z^m (1-z)^{n-m}$, where $G_m^{(n)} \doteq c_{2m}(n!)^2/[(n-m)!(m!)^2]$. Since $z^m(1-z)^{n-m} = z^m - (n-m)z^{m+1} + \cdots$, it follows that the coefficient in z of M_{2n}/M_2^n is $G_1^{(n)} - nG_0^{(n)} = n(n!)c_2 - n(n!)c_0 = 0$, since $c_0 = c_2 = 1$.

B. Computation of the Variance (17)

From the definitions of $t(\hat{\boldsymbol{m}})$, $g(\hat{\boldsymbol{m}})$ and $h_{\mathrm{EOS}}(\hat{\boldsymbol{m}})$, we have

$$\nabla t = \frac{\nabla g}{(1-g)^2} \tag{43}$$

$$\nabla g = \frac{\nabla h_{\mathrm{EOS}}}{F'_{\mathrm{EOS}}(g)} = \frac{\nabla h_{\mathrm{EOS}}}{4F_4 g^3 + 3F_3 g^2 + 2F_2 g}. \tag{44}$$

∇h_{EOS} can be readily computed from (11), obtaining

$$\nabla h_{\mathrm{EOS}} = \frac{1}{\hat{M}_2^5} \hat{\boldsymbol{R}} \boldsymbol{\alpha} \tag{45}$$

$$\hat{\boldsymbol{R}} \doteq \begin{bmatrix} -2\hat{M}_4\hat{M}_2^2 & -3\hat{M}_6\hat{M}_2 & -4\hat{M}_4^2 & -4\hat{M}_8 \\ \hat{M}_2^3 & 0 & 2\hat{M}_4\hat{M}_2 & 0 \\ 0 & \hat{M}_2^2 & 0 & 0 \\ 0 & 0 & 0 & \hat{M}_2 \end{bmatrix}. \tag{46}$$

Now, noting that $g|_{\hat{\boldsymbol{m}}=\boldsymbol{m}} = z = \rho/(1+\rho)$, the evaluation of (43)–(46) at $\hat{\boldsymbol{m}} = \boldsymbol{m}$ yields

$$\boldsymbol{v} = \frac{1}{N^5} \frac{\boldsymbol{R}\boldsymbol{\alpha}}{X(\rho)} \tag{47}$$

where $\boldsymbol{R} \doteq \hat{\boldsymbol{R}}|_{\hat{\boldsymbol{m}}=\boldsymbol{m}}$. The polynomial $X(\rho) \doteq X_3\rho^3 + X_2\rho^2 + X_1\rho$ is linear in $\boldsymbol{\alpha}$: we can write $X_3 \doteq 4F_4 + 3F_3 + 2F_2 = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x}_3$, $X_2 \doteq 3F_3 + 4F_2 = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x}_2$, and $X_1 \doteq 2F_2 = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x}_1$, where the 4×1 vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$ depend on the constellation moments only. Upon defining $\boldsymbol{x}(\rho) \doteq \sum_{n=1}^{3} \boldsymbol{x}_n \rho^n$, one has that $X(\rho) = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x}(\rho)$, and the approximation of the variance can be written as

$$\boldsymbol{v}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{v} = \frac{1}{N^{10}} \frac{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{R}\boldsymbol{\alpha}}{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x}(\rho)\left[\boldsymbol{x}(\rho)\right]^{\mathrm{T}}\boldsymbol{\alpha}}. \tag{48}$$

Observe that $[\boldsymbol{C}]_{ij} = (M_{2(i+j)} - M_{2i}M_{2j})/K$ has the form of (40) for $Q = i + j$, and that the elements of the ith row of $\boldsymbol{R}$ have the form of (40) as well, with $Q = 5 - i$, ($i = 1, 2, 3, 4$). It is then possible to show that all the nonzero elements of $\boldsymbol{R}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{R}$ are of the form of (40) with $Q = 10$. With this, the factor N^{10} in (48) cancels out and (17) follows by simply identifying

$A^{(\mathrm{Var})}(\rho) = K/N^{10}\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{R}\boldsymbol{\alpha}$, and $B^{(\mathrm{Var})}(\rho) = X^2(\rho) = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x}(\rho)[\boldsymbol{x}(\rho)]^{\mathrm{T}}\boldsymbol{\alpha}$, which are polynomials in ρ of degrees 10 and 6, respectively.

C. Proofs of Properties 2–5

That $\boldsymbol{B}^{(\mathrm{Var})}(\rho)$ has rank one (Property 4) is evident from (48), if we identify $\boldsymbol{x}(\rho)[\boldsymbol{x}(\rho)]^{\mathrm{T}}$ with $\boldsymbol{B}^{(\mathrm{Var})}(\rho)$ of (37) [and therefore $\boldsymbol{x}(\rho) = \boldsymbol{b}^{(\mathrm{Var})}(\rho)$]. Similarly, identifying $\boldsymbol{b}_6^{(\mathrm{Var})}$ with $\boldsymbol{x}_3$, so that $\boldsymbol{B}_6^{(\mathrm{Var})} = \boldsymbol{x}_3\boldsymbol{x}_3^{\mathrm{T}}$, then Property 5 follows.

In order to prove that $\boldsymbol{A}^{(\mathrm{Var})}(\rho)$ is rank deficient (Property 2), we note that $\boldsymbol{R}$ has rank 3 since its first and third columns are linearly dependent, see (45). It follows that $\mathrm{rank}(\boldsymbol{A}^{(\mathrm{Var})}(\rho)) = \mathrm{rank}(\boldsymbol{R}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{R}) \leq 3$ for all ρ. The first part of Property 3 immediately follows by noting that $\boldsymbol{A}_{10}^{(\mathrm{Var})} = \lim_{\rho\to\infty} \boldsymbol{A}^{(\mathrm{Var})}(\rho)/\rho^{10}$. Finally, for constellations with $I = 1$, 2 or 3 amplitude levels, it is straightforward (but rather tedious) to explicitly compute $\boldsymbol{A}_{10}^{(\mathrm{Var})}$, as well as to check that its rank does not exceed 0, 1, and 2, respectively.

D. Computation of the Bias (21)

The Hessian matrix of $t(\hat{\boldsymbol{m}})$, $\nabla^2 t$, can be readily computed from (43)–(45) as

$$\begin{aligned}\nabla^2 t &= \frac{1}{(1-g)^2 F'_{\mathrm{EOS}}(g)}\nabla^2 h \\ &\quad + \frac{2F'_{\mathrm{EOS}}(g) - (1-g)F''_{\mathrm{EOS}}(g)}{(1-g)^3\left[F'_{\mathrm{EOS}}(g)\right]^2}\nabla h(\nabla h)^{\mathrm{T}} \quad (49)\\ &= \frac{1}{(1-g)^2 F'_{\mathrm{EOS}}(g)}\frac{1}{\hat{M}_2^6}\hat{\boldsymbol{Y}} + J_t(g)\nabla t(\nabla t)^{\mathrm{T}} \quad (50)\end{aligned}$$

where

$$J_t(g) \doteq \frac{(1-g)\left[2F'_{\mathrm{EOS}}(g) - (1-g)F''_{\mathrm{EOS}}(g)\right]}{F'_{\mathrm{EOS}}(g)} \quad (51)$$

and $\hat{\boldsymbol{Y}}$ is a 4×4 symmetric matrix whose only nonzero elements are given by

$$\begin{aligned}[\hat{\boldsymbol{Y}}]_{11} &\doteq 6\beta\hat{M}_4\hat{M}_2^2 + 12\gamma\hat{M}_6\hat{M}_2 \\ &\quad + 20\left(\delta\hat{M}_4^2 + \epsilon\hat{M}_8\right) \quad (52a)\\ [\hat{\boldsymbol{Y}}]_{12} &= [\hat{\boldsymbol{Y}}]_{21} \doteq -2\beta\hat{M}_2^3 - 8\delta\hat{M}_4\hat{M}_2 \quad (52b)\\ [\hat{\boldsymbol{Y}}]_{13} &= [\hat{\boldsymbol{Y}}]_{31} \doteq -3\gamma\hat{M}_2^2 \quad (52c)\\ [\hat{\boldsymbol{Y}}]_{14} &= [\hat{\boldsymbol{Y}}]_{41} \doteq -4\epsilon\hat{M}_2 \quad (52d)\\ [\hat{\boldsymbol{Y}}]_{22} &\doteq 2\delta\hat{M}_2^2. \quad (52e)\end{aligned}$$

The evaluation of (50) at $\hat{\boldsymbol{m}} = \boldsymbol{m}$ yields

$$\boldsymbol{H} \doteq \nabla^2 t|_{\hat{\boldsymbol{m}}=\boldsymbol{m}} = \frac{1}{N^6}\frac{1}{(1+\rho)X(\rho)}\boldsymbol{Y} + 2J(\rho)\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}} \quad (53)$$

where $\boldsymbol{Y} = \hat{\boldsymbol{Y}}|_{\hat{\boldsymbol{m}}=\boldsymbol{m}}$, $X(\rho)$ has been defined in Section B of this Appendix, and $J(\rho) \doteq (1/2)J_t(\rho/(1+\rho))$. Regarding the first term of (53), its denominator is clearly linear in $\boldsymbol{\alpha}$: $(1+\rho)X(\rho) = \boldsymbol{\alpha}^{\mathrm{T}}[(1+\rho)\boldsymbol{x}(\rho)]$, from where we can identify $\boldsymbol{b}^{(L)}(\rho)$ of (23) as

$$\boldsymbol{b}^{(L)}(\rho) = (1+\rho)\boldsymbol{x}(\rho) = (1+\rho)\boldsymbol{b}^{(\mathrm{Var})}(\rho). \quad (54)$$

Regarding the factor $J(\rho)$ in second term of (53), it is readily seen to be of the form $J(\rho) = A^{(J)}(\rho)/[(1+\rho)X(\rho)] = [\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{a}^{(J)}(\rho)]/[\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{b}^{(L)}(\rho)]$, where $A^{(J)}(\rho) \doteq \sum_{n=0}^{3} A_n^{(J)}\rho^n = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{a}^{(J)}(\rho)$, and $\boldsymbol{a}^{(J)}(\rho)$ depends on ρ and the constellation moments.

Recalling the approximation (20), simple algebra shows that

$$\frac{1}{2}\mathrm{Tr}\{\boldsymbol{H}\boldsymbol{C}\} = \frac{1}{2}\frac{\mathrm{Tr}\{\boldsymbol{Y}\boldsymbol{C}\}}{N^6(1+\rho)X(\rho)} + J(\rho)\boldsymbol{v}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{v} \quad (55)$$

where the second term turns out to be proportional to the approximation for the variance (16). Using (17)–(19) and (54) we can write the second term of (55) as a quotient of cubic forms in $\boldsymbol{\alpha}$

$$J(\rho)\boldsymbol{v}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{v} = \frac{1}{K}\frac{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{a}^{(\mathrm{J})}(\rho)\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}^{(\mathrm{Var})}(\rho)\boldsymbol{\alpha}}{(1+\rho)\left[\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{b}^{(\mathrm{Var})}(\rho)\right]^3}. \quad (56)$$

As for the computation of the first term of (55), which amounts to $L(\rho)$ in (21a), observe that the nonzero $[\boldsymbol{Y}]_{ij}$ have the form of (40) with $Q = 6 - (i+j)$. We are eventually interested in the elements of the diagonal of $\boldsymbol{H}\boldsymbol{C}$, which are given by

$$[\boldsymbol{H}\boldsymbol{C}]_{ii} = \frac{\sum_{j=1}^{4}[\boldsymbol{Y}]_{ij}[\boldsymbol{C}]_{ji}}{N^6(1+\rho)X(\rho)} \quad (57)$$

for $i = 1, 2, 3, 4$. All terms $([\boldsymbol{Y}]_{ij}[\boldsymbol{C}]_{ji})$ have then the form of (40) with $Q = 6$. Clearly, N^6 cancels out in (57), so that the first term in (21) follows if we identify $L(\rho) = A^{(L)}(\rho)/B^{(L)}(\rho)$, where $A^{(L)}(\rho) = K\sum_{i=1}^{4}\sum_{j=1}^{4}[\boldsymbol{Y}]_{ij}[\boldsymbol{C}]_{ji}/(2N^6)$ and $B^{(L)}(\rho) = (1+\rho)X(\rho)$. Besides, $A^{(L)}(\rho)$ can be readily seen to be linear in $\boldsymbol{\alpha}$, since it is a linear combination of the elements of $\boldsymbol{Y}$, which are themselves linear in $\boldsymbol{\alpha}$ [cf. (50)]; this allows to write $A^{(L)}(\rho) = \sum_{n=0}^{6} A_n^{(L)}\rho^n = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{a}^{(L)}(\rho)$ as in (22), for some $\boldsymbol{a}^{(L)}(\rho) = \sum_{n=0}^{6}\boldsymbol{a}_n^{(L)}\rho^n$, in which $\boldsymbol{a}_n^{(L)}$ depends only on the constellation moments for all n.

ACKNOWLEDGMENT

The authors would like to thank Dr. W. Gappmair for generously sharing his knowledge on numerical computation of CRBs.

REFERENCES

[1] K. Balachandran, S. R. Kadaba, and S. Nanda, "Channel quality estimation and rate adaptation for cellular mobile radio," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 7, pp. 1244–1256, Jul. 1999.

[2] G. Albertazzi, S. Cioni, G. Corazza, M. Neri, R. Pedone, P. Salmi, A. Vanelli-Coralli, and M. Villanti, "On the adaptive DVB-S2 physical layer: Design and performance," *IEEE Wireless Commun. Mag.*, vol. 12, no. 6, pp. 62–68, Dec. 2005.

[3] T. A. Summers and S. G. Wilson, "SNR mismatch and online estimation in turbo decoding," *IEEE Trans. Commun.*, vol. 46, no. 4, pp. 421–423, Apr. 1998.

[4] D. Pauluzzi and N. Beaulieu, "A comparison of SNR estimation techniques for the AWGN channel," *IEEE Trans. Commun.*, vol. 48, no. 10, pp. 1681–1691, Oct. 2000.

[5] J. F. Sevillano, I. Vélez, M. Leyh, S. Lipp, A. Irizar, and L. Fontán, "In-service SNR estimation without symbol timing recovery for QPSK data transmission systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 9, pp. 3202–3207, Sep. 2007.

[6] R. B. Kerr, "On signal and noise level estimation in a coherent PCM channel," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-2, no. 4, pp. 450–454, 1966.

[7] R. Gagliardi and C. Thomas, "PCM data reliability monitoring through estimation of signal-to-noise ratio," *IEEE Trans. Commun.*, vol. 16, no. 3, pp. 479–486, Jun. 1968.
[8] T. Benedict and T. Soong, "The joint estimation of signal and noise from the sum envelope," *IEEE Trans. Inf. Theory*, vol. 13, no. 3, pp. 447–454, Jul. 1967.
[9] R. Matzner, "An SNR estimation algorithm for complex baseband signals using higher order statistics," *Facta Universitatis (Niš)*, vol. 6, no. 1, pp. 41–52, 1993.
[10] P. Gao and C. Tepedelenlioğlu, "SNR estimation for nonconstant modulus constellations," *IEEE Trans. Signal Process.*, vol. 53, no. 3, pp. 865–870, Mar. 2005.
[11] A. Wiesel, J. Goldberg, and H. Messer, "Non-data-aided signal-to-noise-ratio estimation," in *Proc. IEEE Int. Conf. Commun. ICC 2002*, New York City, Apr. 28–May 2 2002, vol. 1, pp. 197–201.
[12] A. Das, "NDA SNR estimation: CRLBs and EM based estimators," in *Proc. IEEE Region 10 Conf. TENCON*, Hyderabad, India, Nov. 18–21, 2008, pp. 1–6.
[13] R. López-Valcarce, C. Mosquera, and W. Gappmair, "Iterative envelope-based SNR estimation for nonconstant modulus constellations," in *Proc. IEEE 8th Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Helsinki, Finland, Jun. 17–20, 2007, pp. 1–5.
[14] R. López-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.
[15] W. Gappmair and O. Koudelka, "Moment-based SNR estimation of signals with non-constant envelope," in *Proc. 3rd Adv. Satellite Mobile Syst. Conf.*, Herrsching, Germany, May 2006, pp. 301–303.
[16] W. Gappmair, R. López-Valcarce, and C. Mosquera, "Cramer-Rao lower bound and EM algorithm for envelope-based SNR estimation of nonconstant modulus constellations," *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1622–1627, Jun. 2009.
[17] M. Türkboylari and G. Stüber, "An efficient algorithm for estimating the signal-to-interference ratio in TDMA cellular systems," *IEEE Trans. Commun.*, vol. 46, no. 6, pp. 728–731, Jun. 1998.
[18] A. Wiesel, J. Goldberg, and H. Messer-Yaron, "SNR estimation in time-varying fading channels," *IEEE Trans. Commun.*, vol. 54, no. 5, pp. 841–848, May 2006.
[19] A. Stephenne, F. Bellili, and S. Affes, "Moment-based SNR estimation for SIMO wireless communication systems using arbitrary QAM," in *Conf. Rec. 41st Asilomar Conf. Signals, Syst. Comput. (ACSSC)*, Pacific Grove, CA, Nov. 4–7, 2007, pp. 601–605.
[20] F. Bellili, A. Stephenne, and S. Affes, "SNR estimation of QAM-modulated transmissions over time-varying SIMO channels," in *Proc. IEEE Int. Symp. Wireless Commun. Syst. (ISWCS)*, Reykjavik, Iceland, Oct. 21–24, 2008, pp. 199–203.
[21] M. A. Dangl and J. Lindner, "How to use a priori information of data symbols for SNR estimation," *IEEE Signal Process. Lett.*, vol. 13, no. 11, pp. 661–664, Nov. 2006.
[22] G. Hatzichristos and M. P. Fargues, "A hierarchical approach to the classification of digital modulation types in multipath environments," in *Conf. Rec. 35th Asilomar Conf. Signals, Syst. Comput.*, Nov. 4–7, 2001, vol. 2, pp. 1494–1498.
[23] W. Dai, Y. Wang, and J. Wang, "Joint power estimation and modulation classification using second- and higher statistics," in *Proc. IEEE WCNC2002 Wireless Commun. Network. Conf.*, Mar. 17–21, 2002, vol. 1, pp. 155–158.
[24] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of automatic modulation classification techniques: Classical approaches and new trends," *IET Commun.*, vol. 1, no. 2, pp. 137–156, Apr. 2007.
[25] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
[26] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.
[27] C.-Y. Huang and A. Polydoros, "Likelihood methods for MPSK modulation classification," *IEEE Trans. Commun.*, vol. 43, no. 234, pp. 1493–1504, Feb.–Mar.–Apr. 1995.
[28] *Handbook of Mathematical Functions*, M. Abramowitz and I. Stegun, Eds., 9th ed. New York: Dover, 1972.
[29] *Digital Video Broadcasting (DVB); Second Generation Framing Structure, Channel Coding and Modulation Systems for Broadcasting, Interactive Services, News Gathering and Other Broadband Satellite Applications*, EN 300 307 V1.1.2, ETSI, Jun. 2006.
[30] Signal Processing in Communications Group, Univ. Vigo, Vigo, Spain, "Mathematica and MATLAB code supporting some of the derivations in the paper," 2009. [Online]. Available: http://www.gts.tsc.uvigo.es/~valcarce/snr_hos.html

Marcos Álvarez-Díaz (S'05) received the M.S. degree in telecommunications engineering from the University of Vigo, Vigo, Spain, in 2002, where he is currently pursuing the Ph.D. degree.

During 2003, he was with the Institute of Communications and Navigation, German Aerospace Center (DLR), Oberpfaffenhofen, Germany. Since 2004, he has been with Signal Processing, Communications Group, Signal Theory and Communications Department, University of Vigo. His research interests include statistical signal processing, parameter estimation and synchronization.

Roberto López-Valcarce (M'01) received the undergraduate diploma in telecommunication engineering from Universidad de Vigo, Vigo, Spain, in 1995, and the M.S. and Ph.D. degrees in electrical engineering from the University of Iowa, Iowa City, in 1998 and 2000, respectively.

During 1995, he was a Project Engineer with Intelsis. He was a Postdoctoral Fellow of the Spanish Ministry of Science and Technology from 2001 to 2006. During that period, he was with the Signal Theory and Communications Department, Universidad de Vigo, where he is currently an Associate Professor. His main research interests lie in the area of adaptive signal processing, digital communications, and cognitive radio. He has coauthored over 20 journal and 40 conference papers on these topics. He holds two European patents in the field of adaptive processing applied to communications systems.

Dr. López-Valcarce was a recipient of a 2005 Best Paper Award of the IEEE Signal Processing Society and a coauthor of the textbook "Comunicaciones Digitales" (Prentice Hall, 2007) (in Spanish). Since 2008 he serves as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING.

Carlos Mosquera (S'93–M'98) received the M.Sc. degree from Stanford University, Stanford, CA, in 1994, and the Ph.D. degree from University of Vigo, Vigo, Spain, in 1998, both in electrical engineering.

He has been an Associate Professor with the University of Vigo since 2000. He held visiting positions with the European Space Agency (1999) and the University of New Mexico (2007), respectively. He holds four patents in collaboration with industry, and has coauthored over 20 journal papers and two books. He has been the principal investigator of a number of national and European projects and research contracts, and has been member of several international technical program committees. He is currently the Research director of Gradiant (Galician Research and Development Center in Advanced Telecommunications) in digital communications.