

On the Use of Extrinsic Probabilities in the Computation of non-Bayesian Cramér-Rao Bounds for Coded Linearly Modulated Signals

Nele Noels and Marc Moeneclaey
Ghent University, Belgium
{Nele.Noels, Marc.Moeneclaey}@UGent.be

Abstract—This contribution considers the non-Bayesian Cramér-Rao bound (CRB) related to parameter estimation from a linearly modulated signal observed in additive white Gaussian noise. We compare the exact CRB expression for coded modulation with two ad hoc CRB (ACRB) expressions; the first ACRB is obtained by substituting in the exact CRB expression for uncoded modulation the a priori symbol probabilities by the extrinsic symbol probabilities; the second ACRB, which has received some attention in recent scientific publications, additionally assumes that the real and imaginary parts of the symbols are independent. Our exposition focuses on the particular case of phase shift estimation. By means of examples we show that although for some coded modulation schemes the exact and ad hoc CRBs yield virtually the same numerical result, for other coded modulation schemes the ACRBs differ considerably among themselves and from the exact CRB. We provide some explanations for this behavior. We also argue that both ACRBs are expected to virtually coincide with the exact CRB in the case of bit-interleaved coded modulation (BICM) combined with a rectangular constellations with independent in-phase and quadrature mapping and a binary code for which the factor graph does not contain short cycles.

I. INTRODUCTION

The synchronization subsystems of a digital communication receiver have the purpose of achieving the correct alignment of the incoming waveform with certain locally generated reference signals [1]. For instance, a coherent receiver for bandpass transmission needs carrier phase synchronization. This means that the difference between the phase of the local sinusoid that is used for bandpass-to-baseband conversion and that of the incoming carrier (the phase shift) must be estimated and compensated for.

A common approach to assess the performance of a synchronization parameter estimator in a digital receiver is to compare the resulting mean square error (MSE) to the Cramér-Rao Bound (CRB). We consider the unknown parameter to be estimated as deterministic; hence, the relevant CRB is the non-Bayesian CRB, which is a fundamental lower bound on the error variance of unbiased estimators [2]–[4]. In many cases, the statistics of the observation depend not only on the parameter to be estimated, but also on a set of random nuisance parameters with a known distribution, independent of the parameter we want to estimate. The presence of these random nuisance parameters can make the analytical computation of the CRB very hard, if not impossible.

In order to avoid the computational complexity associated with the true CRB, a modified CRB (MCRB) has been derived

in [5]. The MCRB is much simpler to evaluate than the CRB, but is in general looser than the CRB. In [6], the high Signal-to-Noise Ratio (SNR) limit of the CRB, pertaining to the estimation of a deterministic scalar parameter, has been evaluated analytically, and has been shown to coincide with the MCRB when nuisance parameters are discrete-valued. Assuming statistically independent and uniformly distributed data symbols (which holds for uncoded linear modulation), several algorithms to efficiently evaluate true CRBs with a reasonable complexity have been proposed, see, e.g., [7]–[10] for CRBs corresponding to phase estimation. In [11], the low-SNR limit of these uncoded CRBs has been obtained analytically.

In the presence of coding, synchronization algorithms must operate at a low SNR, so that the high-SNR limit of the CRB might no longer be a relevant benchmark. Moreover, the code structure introduces statistical dependence among the data symbols. This has two important consequences. First, the CRBs derived for statistically independent and uniformly distributed data symbols no longer apply. Second, evaluating valid CRBs is significantly more involved for coded than for uncoded transmissions.

Numerical procedures enabling the efficient evaluation of (exact or approximate) CRB expressions for synchronization parameter estimation in the presence of coded linear modulation have been presented in [12]–[20]. In [12]–[15], an *exact* expression of the CRB is derived in terms of the marginal symbol a posteriori probabilities. In [17]–[20], for rectangular constellations with independent mapping on the real and imaginary part of the symbols, approximate CRB expressions are derived under the simplifying assumption that the transmitted data symbols are statistically independent, with the marginal a priori symbol probabilities substituted by the corresponding extrinsic probabilities. Although this *ad hoc* approach lacks mathematical justification, the numerical results presented in [17]–[20] for bit-interleaved turbo-coded Gray-mapped square quadrature amplitude modulations (QAM) turn out to be rather close to the ones obtained using the correct expression (as in [12]–[15]). The present paper (i) investigates the validity of the ad hoc CRB (ACRB) expressions from [17]–[20]; (ii) compares for various coded modulation systems the numerical values resulting from the ACRB expressions and the exact CRB expression from [12]–[15]; and (iii) assesses the computational complexity associated with the numerical

evaluation of the CRB and ACRB expressions. For ease of exposition we focus on the simple case of phase estimation.

The outline is as follows. Section II provides the observation model and the corresponding CRB definition. Section III reviews the correct CRB expression for coded modulation systems, presented in [12]. In Section IV we derive a simpler CRB expression that holds for independent data symbols only. Section V builds upon the result from Section IV to obtain two ad hoc CRB expressions for coded modulation systems, one of which is used in [18]. Sufficient conditions for these ACRBs to equal the exact CRB are formulated in Section VI. Section VII discusses the complexity associated with the numerical evaluation of the various CRB expressions. Numerical results pertaining to various coded modulation systems are presented in Section VIII. Final conclusions are drawn in Section IX. The main contributions of the presented work are the following: (i) using arguments based on factor graphs, we explain why both ACRBs essentially coincide with the CRB in the case of BICM using codes without short cycles and square QAM modulation with independent in-phase and quadrature mapping; and (ii) we point out that, depending on the specific coded modulation system considered, one or both of the ACRBs can substantially differ from the CRB, which implies that the ACRBs should be used with great caution.

Throughout this paper we will adopt the following notations and conventions:

- K : number of observed data symbols
- $\mathbf{v} = (v_1, v_2, \dots, v_K)$: row vector of size K with components v_k
- $\mathbf{v}_R = \text{Re}(\mathbf{v})$ and $\mathbf{v}_I = \text{Im}(\mathbf{v})$: real and imaginary parts of \mathbf{v}
- $v_{R,k}$ and $v_{I,k}$: real and imaginary parts of v_k
- $\mathbf{v}^{(k)} = (v_1, v_2, \dots, v_{k-1}, v_{k+1}, \dots, v_K)$: row vector obtained by removing v_k from \mathbf{v}
- $\mathbf{v}^{(k,R)} = (\mathbf{v}^{(k)}, v_{I,k})$: row vector obtained by removing $v_{R,k}$ from \mathbf{v}
- $\mathbf{v}^{(k,I)} = (\mathbf{v}^{(k)}, v_{R,k})$: row vector obtained by removing $v_{I,k}$ from \mathbf{v}
- $\mathbb{E}_{\mathbf{u}} [f(\mathbf{u})]$: expectation of $f(\mathbf{u})$ over \mathbf{u}
- $\mathbb{E}_{\mathbf{u}|\mathbf{z}} [f(\mathbf{u})]$: expectation of $f(\mathbf{u})$ over \mathbf{u} , conditioned on \mathbf{z}
- $\mathbb{E}[f(\mathbf{u})|\mathbf{z}]$: short-hand notation for $\mathbb{E}_{\mathbf{u}|\mathbf{z}} [f(\mathbf{u})]$
- $p(\mathbf{u}|\mathbf{v}) \propto g(\mathbf{u}, \mathbf{v})$: the conditional probability mass function (pmf) $p(\mathbf{u}|\mathbf{v})$ equals $C(\mathbf{v})g(\mathbf{u}, \mathbf{v})$, with $C(\mathbf{v})$ a normalization factor not depending on \mathbf{u}
- $\mathbf{1} = (1, 1, \dots, 1)^T$: all-ones column vector of size K
- $\text{diag}(\mathbf{V}) = (V_{1,1}, V_{2,2}, \dots, V_{K,K})$: row vector containing the diagonal of a square $K \times K$ matrix \mathbf{V} with elements $V_{k,l}$.

II. OBSERVATION MODEL AND CRAMÉR-RAO BOUND

We consider the transmission of an information bit sequence of length N_b over an AWGN channel using coded linear modulation. As the receiver has no prior knowledge about the transmitted information bits, these bits are modeled as independent binary random variables, uniformly distributed over the set $\{0, 1\}$. The information bits are applied to a binary encoder with rate r_c . The resulting N_b/r_c coded bits are mapped to a data symbol sequence $\mathbf{x} = (x_1, x_2, \dots, x_K)$,

where the symbols x_k belong to a unit-energy complex-valued M -point constellation, and $K = N_b / (r_c \log_2(M))$ is the length of the symbol sequence. The probability mass function (pmf) $p(\mathbf{x})$ of \mathbf{x} equals 2^{-N_b} if \mathbf{x} belongs to the set $\mathbb{S}_{\mathbf{x}}$ of legitimate coded symbol sequences, and 0 otherwise. The receiver observes the vector \mathbf{y} of size K , given by

$$\mathbf{y} = \mathbf{x}e^{j\theta} + \mathbf{w}. \quad (1)$$

The components w_k of \mathbf{w} are independent and identically distributed, circularly symmetric complex Gaussian random variables with zero mean and variance σ^2 . The deterministic parameter θ represents the unknown phase shift introduced by the channel. Coherent receivers need to extract from the observation \mathbf{y} an estimate $\hat{\theta}(\mathbf{y})$ of the phase shift θ , to correct the received signal before decoding the received information. Consequently, the accuracy of $\hat{\theta}(\mathbf{y})$ has direct repercussions on the error rate of the decoder.

Estimation accuracy is usually expressed in terms of the MSE of the estimate¹. Fundamental bounds on the achievable estimator MSE serve as a useful benchmark for assessing the performance of practical estimators. One of the most celebrated performance limits is the CRB; for unbiased $\hat{\theta}(\mathbf{y})$, we have [4]:

$$\text{MSE}(\theta) \geq I^{-1}(\theta), \quad (2)$$

where $\text{MSE}(\theta) = \mathbb{E}_{\mathbf{y}} \left[\left(\hat{\theta}(\mathbf{y}) - \theta \right)^2 \right]$, and

$$I(\theta) = \mathbb{E}_{\mathbf{y}} \left[\left(\frac{\partial}{\partial \theta} \ln p(\mathbf{y}; \theta) \right)^2 \right] \quad (3)$$

denotes the Fisher Information (FI), and $p(\mathbf{y}; \theta)$ is the probability density function (pdf) of \mathbf{y} corresponding to a phase shift equal to θ . It follows immediately from (1) that $p(\mathbf{y}; \theta)$ required in (3) can be computed as

$$p(\mathbf{y}; \theta) = \mathbb{E}_{\mathbf{x}} [p(\mathbf{y}|\mathbf{x}; \theta)], \quad (4)$$

with

$$p(\mathbf{y}|\mathbf{x}; \theta) = \prod_{k=1}^K p(y_k | x_k; \theta), \quad (5)$$

where

$$p(y_k | x_k; \theta) = \frac{1}{\pi\sigma^2} e^{-\frac{|y_k - x_k e^{j\theta}|^2}{\sigma^2}}. \quad (6)$$

Conceptually, the evaluation of the CRB (2)-(3) is straightforward. However, the difficulty lies in the fact that the statistics of the observed vector \mathbf{y} depend not only on the phase shift θ to be estimated, but also on the unknown coded

¹As the observation (1) is periodic in θ (with period 2π), the appropriate performance measure is the *periodic* MSE (PMSE), defined as $\text{PMSE}(\theta) = \mathbb{E}_{\mathbf{y}} \left[\left([\hat{\theta}(\mathbf{y}) - \theta]_{2\pi} \right)^2 \right]$, where $[u]_{2\pi}$ denotes the modulo- 2π reduction of u to the interval $[-\pi, \pi)$. As outlined in [21], the PMSE related to a periodic-unbiased estimator is lower bounded by the periodic CRB (PCRb), which differs from the standard CRB. However, under practical operating conditions (where the PMSE is sufficiently small to cause only a minor degradation of the decoder performance after phase correction), the PCRb virtually coincides with the standard CRB (which lower bounds the standard MSE of an unbiased estimate), while the latter is easier to evaluate. Therefore, the vast majority of the literature dealing with performance bounds on carrier phase and symbol timing estimation considers the CRB rather than the PCRb.

data symbol sequence \mathbf{x} . As a result, the evaluation of $p(\mathbf{y}; \theta)$ from (4) requires a sum ranging over all 2^{N_b} legitimate symbol vectors \mathbf{x} in \mathbb{S}_x . In general, the resulting computational burden is exponential in N_b , making the computation of the CRB potentially very hard, if not impossible.

In order to avoid the computational complexity associated with the summation over \mathbf{x} in (4), a simpler but looser bound, referred to as the MCRB, has been derived [5]:

$$\text{MSE}(\theta) \geq I_m^{-1}(\theta), \quad (7)$$

where

$$I_m(\theta) = \mathbb{E}_{\mathbf{y}, \mathbf{x}} \left[\left(\frac{\partial}{\partial \theta} \ln p(\mathbf{y}, \mathbf{x}; \theta) \right)^2 \right] \quad (8)$$

is the modified FI, which satisfies $I^{-1}(\theta) \geq I_m^{-1}(\theta)$. As opposed to the CRB, the MCRB is expressed directly in terms of the joint pdf $p(\mathbf{y}, \mathbf{x}; \theta) = p(\mathbf{y} | \mathbf{x}; \theta) p(\mathbf{x})$, rather than the more complicated $p(\mathbf{y}; \theta)$. For the case at hand, the MCRB (7) simply yields $\text{MSE}(\theta) \geq \sigma^2 / (2K)$, which does not depend on the underlying coded modulation scheme.

Noting from (4)-(6) that $p(\mathbf{y}; \theta) = p(\mathbf{y} e^{j\psi}; \theta + \psi)$, for any angle ψ , it follows that $I(\theta)$ is independent of θ ; in the sequel we consider $\theta = 0$, and introduce the concise notations $p(\mathbf{y} | \mathbf{x})$, $p(\mathbf{y})$ and $p(\mathbf{x} | \mathbf{y})$ for $p(\mathbf{y} | \mathbf{x}, \theta = 0)$, $p(\mathbf{y} | \theta = 0)$ and $p(\mathbf{x} | \mathbf{y}, \theta = 0)$, respectively.

III. CORRECT CRB EXPRESSION

A correct expression for the FI (3) related to the observation model (1) has been derived in [12], where $\frac{\partial}{\partial \theta} \ln p(\mathbf{y}; \theta)$ has been expressed in terms of the marginal symbol *a posteriori* probabilities $p(x_k | \mathbf{y}; \theta)$. Considering $\theta = 0$ without loss of generality, [12] yields:

$$\frac{\partial}{\partial \theta} \ln p(\mathbf{y}; \theta) \Big|_{\theta=0} = \frac{2}{\sigma^2} (\mathbf{a}(\mathbf{y}) \cdot \mathbf{1}), \quad (9)$$

where $\mathbf{a}(\mathbf{y}) = (a_1(\mathbf{y}), a_2(\mathbf{y}), \dots, a_K(\mathbf{y}))$, with

$$a_k(\mathbf{y}) = y_{I,k} \mathbb{E}[x_{R,k} | \mathbf{y}] - y_{R,k} \mathbb{E}[x_{I,k} | \mathbf{y}]. \quad (10)$$

No approximations are involved in obtaining (9). Substituting (9) into (3) yields:

$$I(0) = \frac{4}{\sigma^4} (\mathbf{1}^T \cdot \mathbf{D} \cdot \mathbf{1}), \quad (11)$$

where $\mathbf{D} = \mathbb{E}_{\mathbf{y}} [\mathbf{a}^T(\mathbf{y}) \mathbf{a}(\mathbf{y})]$ denotes the covariance matrix of $\mathbf{a}(\mathbf{y})$. The (k, l) th element of \mathbf{D} is given by

$$D_{k,l} = \mathbb{E}_{\mathbf{y}} [a_k(\mathbf{y}) a_l(\mathbf{y})]. \quad (12)$$

IV. CORRECT CRB EXPRESSION FOR INDEPENDENT SYMBOLS

In the case of statistically independent data symbols (which holds for uncoded transmission but not in the presence of channel coding²), the computation of $I(0)$ simplifies significantly. Indeed, assuming $p(\mathbf{x}) = \prod_k p(x_k)$, it follows from (5) that $p(\mathbf{y}) = \prod_k p(y_k)$ and $p(\mathbf{x} | \mathbf{y}) = \prod_k p(x_k | y_k)$,

²We exclude ‘‘pathological’’ cases, such as all N_b/r_c bits from a code-word being mapped to a single data symbol (which corresponds to $N_b = r_c \log_2(M)$), yielding $K = 1$).

so that $p(x_k | \mathbf{y}) = p(x_k | y_k)$ and, hence, $a_k(\mathbf{y})$ from (10) depends on y_k only: $a_k(\mathbf{y}) = \check{a}_k(y_k)$ with

$$\check{a}_k(y_k) = y_{I,k} \mathbb{E}[x_{R,k} | y_k] - y_{R,k} \mathbb{E}[x_{I,k} | y_k]. \quad (13)$$

For $k \neq l$, the quantity $D_{k,l}$ from (12) now decomposes as $D_{k,l} = \mathbb{E}_{y_k} [\check{a}_k(y_k)] \mathbb{E}_{y_l} [\check{a}_l(y_l)]$, where

$$\begin{aligned} & \mathbb{E}_{y_k} [\check{a}_k(y_k)] \\ &= \mathbb{E}_{y_k} [y_{I,k} \mathbb{E}[x_{R,k} | y_k] - y_{R,k} \mathbb{E}[x_{I,k} | y_k]] \\ &= \mathbb{E}_{x_{R,k}, x_{I,k}} [x_{R,k} \mathbb{E}[y_{I,k} | x_{I,k}] - x_{I,k} \mathbb{E}[y_{R,k} | x_{R,k}]] \\ &= \mathbb{E}_{x_k} [x_{R,k} x_{I,k} - x_{I,k} x_{R,k}] = 0, \end{aligned}$$

so that in (11) \mathbf{D} becomes a diagonal matrix with the k th diagonal element given by

$$D_{k,k} = \mathbb{E}_{y_k} [(\check{a}_k(y_k))^2]. \quad (14)$$

We obtain:

$$I(0) = \frac{4}{\sigma^4} (\text{diag}(\mathbf{D}) \cdot \mathbf{1}). \quad (15)$$

Note that the evaluation of $D_{k,k}$ requires the distributions $p(x_{R,k} | y_k)$, $p(x_{I,k} | y_k)$ (in (13)) and $p(y_k)$ (in (14)), which can be computed from $p(y_k | x_k)$ and $p(x_k)$ by applying Bayes’ rule.

Let us consider the stronger assumption, that the real and imaginary parts of the symbols are statistically independent, i.e., $p(\mathbf{x}) = \prod_k p(x_{R,k}) p(x_{I,k})$; this assumption holds for uncoded transmission of QAM symbols with independent mapping on $x_{R,k}$ and $x_{I,k}$, but not for coded modulation schemes. In this case, we obtain $p(x_{R,k}, x_{I,k}) = p(x_{R,k}) p(x_{I,k})$, which in turn yields:

$$p(x_{R,k}, x_{I,k} | y_{R,k}, y_{I,k}) = p(x_{R,k} | y_{R,k}) p(x_{I,k} | y_{I,k}) \quad (16)$$

and

$$p(y_{R,k}, y_{I,k}) = p(y_{R,k}) p(y_{I,k}). \quad (17)$$

Using (16) and (17), $D_{k,k}$ from (14) transforms into

$$D_{k,k} = \mathbb{E}_{y_{R,k}} \left[\mathbb{E}_{y_{I,k}} [(\check{a}_k(y_k))^2] \right], \quad (18)$$

where $\check{a}_k(y_k)$ simplifies to

$$\check{a}_k(y_k) = y_{I,k} \mathbb{E}[x_{R,k} | y_{R,k}] - y_{R,k} \mathbb{E}[x_{I,k} | y_{I,k}]. \quad (19)$$

Note the subtle differences between (13)-(14) and (18)-(19). The distributions $p(x_{R,k} | y_{R,k})$, $p(x_{I,k} | y_{I,k})$, $p(y_{R,k})$ and $p(y_{I,k})$ needed in (18)-(19) can be obtained from $p(y_{R,k} | x_{R,k})$, $p(y_{I,k} | x_{I,k})$, $p(x_{R,k})$ and $p(x_{I,k})$.

V. AD HOC CRB EXPRESSIONS

We consider two ad hoc CRB (ACRB) expressions for coded modulation. The ACRB expressions are inspired on the CRB for uncoded transmission, derived in Section IV. The first ACRB expression is obtained from (13)-(15), with the a priori symbol probabilities simply replaced by the extrinsic symbol probabilities. The second ACRB expression (which has been used in [18]) results from (15) and (18)-(19), with the a priori probabilities of the real and imaginary parts of the data symbols substituted by the corresponding extrinsic probabilities.

A. First ACRB Expression

The first ACRB expression is obtained as follows. First, use (13)-(15), where in the computation of the distributions $p(x_k | y_k)$ and $p(y_k)$ the a priori pmf $p(x_k)$ is replaced by the extrinsic pmf $p(x_k | \mathbf{y}^{(k)} = \tilde{\mathbf{y}}^{(k)})$ corresponding to a specific realization $\tilde{\mathbf{y}}$ of the received signal vector; we denote the result as $I^{(1)}(0; \tilde{\mathbf{y}})$. Next, compute the average $I^{(1)}(0) = \mathbb{E}_{\tilde{\mathbf{y}}} [I^{(1)}(0; \tilde{\mathbf{y}})]$. The corresponding ACRB is given by $1/I^{(1)}(0)$, and will be denoted $\text{ACRB}^{(1)}$.

Taking into account Bayes' rule, it is easily verified that replacing the a priori symbol pmfs by the extrinsic symbol pmfs corresponds to substituting in (13)-(15) $p(x_k | y_k)$ and $p(y_k)$ by $p(x_k | \mathbf{y})|_{\mathbf{y}^{(k)} = \tilde{\mathbf{y}}^{(k)}}$ and $p(y_k | \mathbf{y}^{(k)} = \tilde{\mathbf{y}}^{(k)})$, respectively. This yields

$$I^{(1)}(0; \tilde{\mathbf{y}}) = \frac{4}{\sigma^4} (\mathbf{d}^{(1)}(\tilde{\mathbf{y}}) \cdot \mathbf{1}),$$

with $\mathbf{d}^{(1)}(\tilde{\mathbf{y}}) = (d_1^{(1)}(\tilde{\mathbf{y}}^{(1)}), \dots, d_K^{(1)}(\tilde{\mathbf{y}}^{(K)}))$, where

$$d_k^{(1)}(\tilde{\mathbf{y}}^{(k)}) = \mathbb{E}_{y_k | \tilde{\mathbf{y}}^{(k)}} \left[\left(a_k^{(1)}(y_k, \tilde{\mathbf{y}}^{(k)}) \right)^2 \right] \quad (20)$$

and

$$a_k^{(1)}(y_k, \tilde{\mathbf{y}}^{(k)}) = y_{I,k} \mathbb{E}[x_{R,k} | \mathbf{y}]|_{\mathbf{y}^{(k)} = \tilde{\mathbf{y}}^{(k)}} - y_{R,k} \mathbb{E}[x_{I,k} | \mathbf{y}]|_{\mathbf{y}^{(k)} = \tilde{\mathbf{y}}^{(k)}}. \quad (21)$$

The average $I^{(1)}(0) = \mathbb{E}_{\tilde{\mathbf{y}}} [I^{(1)}(0; \tilde{\mathbf{y}})]$ can be represented as

$$I^{(1)}(0) = \frac{4}{\sigma^4} (\bar{\mathbf{d}}^{(1)} \cdot \mathbf{1}), \quad (22)$$

with $\bar{\mathbf{d}}^{(1)} = (\bar{d}_1^{(1)}, \bar{d}_2^{(1)}, \dots, \bar{d}_K^{(1)})$ and

$$\bar{d}_k^{(1)} = \mathbb{E}_{y_k, \tilde{\mathbf{y}}^{(k)}} \left[\left(a_k^{(1)}(y_k, \tilde{\mathbf{y}}^{(k)}) \right)^2 \right].$$

Noting from (10) and (21) that $a_k^{(1)}(y_k, \tilde{\mathbf{y}}^{(k)}) = a_k(\mathbf{y})|_{\mathbf{y}^{(k)} = \tilde{\mathbf{y}}^{(k)}}$, (22) reduces to

$$I^{(1)}(0) = \frac{4}{\sigma^4} (\text{diag}(\mathbf{D}) \cdot \mathbf{1}), \quad (23)$$

where $\mathbf{D} = \mathbb{E}_{\mathbf{y}} [\mathbf{a}^T(\mathbf{y}) \mathbf{a}(\mathbf{y})]$.

B. Second ACRB Expression

The second ACRB expression makes use of (15), and (18)-(19), with replacing the a priori pmfs $p(x_{R,k})$ and $p(x_{I,k})$ by the extrinsic pmfs $p(x_{R,k} | \mathbf{y}^{(k,R)} = \tilde{\mathbf{y}}^{(k,R)})$ and $p(x_{I,k} | \mathbf{y}^{(k,I)} = \tilde{\mathbf{y}}^{(k,I)})$, corresponding to a specific realization $\tilde{\mathbf{y}}$ of the received signal vector. Next, the resulting $I^{(2)}(0; \tilde{\mathbf{y}})$ is averaged over $p(\tilde{\mathbf{y}})$, yielding $I^{(2)}(0)$. The corresponding ad hoc CRB is denoted $\text{ACRB}^{(2)}$, with $\text{ACRB}^{(2)} = 1/I^{(2)}(0)$.

Using Bayes' rule, it is easily shown that replacing the a priori pmfs of $x_{R,k}$ and $x_{I,k}$ by the corresponding extrinsic pmfs is equivalent to substituting in (18)-(19) $p(x_{R,k} | y_{R,k})$, $p(x_{I,k} | y_{I,k})$, $p(y_{R,k})$ and $p(y_{I,k})$ by $p(x_{R,k} | \mathbf{y})|_{\mathbf{y}^{(k,R)} = \tilde{\mathbf{y}}^{(k,R)}}$, $p(x_{I,k} | \mathbf{y})|_{\mathbf{y}^{(k,I)} = \tilde{\mathbf{y}}^{(k,I)}}$,

$p(y_{R,k} | \mathbf{y}^{(k,R)} = \tilde{\mathbf{y}}^{(k,R)})$ and $p(y_{I,k} | \mathbf{y}^{(k,I)} = \tilde{\mathbf{y}}^{(k,I)})$, respectively. This yields:

$$I^{(2)}(0; \tilde{\mathbf{y}}) = \frac{4}{\sigma^4} (\mathbf{d}^{(2)}(\tilde{\mathbf{y}}) \cdot \mathbf{1}),$$

where $\mathbf{d}^{(2)}(\tilde{\mathbf{y}}) = (d_1^{(2)}(\tilde{\mathbf{y}}), d_2^{(2)}(\tilde{\mathbf{y}}), \dots, d_K^{(2)}(\tilde{\mathbf{y}}))$, with

$$d_k^{(2)}(\tilde{\mathbf{y}}) = \mathbb{E}_{y_{R,k} | \tilde{\mathbf{y}}^{(k,R)}} \left[\mathbb{E}_{y_{I,k} | \tilde{\mathbf{y}}^{(k,I)}} \left[\left(a_k^{(2)}(y_k, \tilde{\mathbf{y}}) \right)^2 \right] \right], \quad (24)$$

and

$$a_k^{(2)}(y_k, \tilde{\mathbf{y}}) = y_{I,k} \mathbb{E}[x_{R,k} | \mathbf{y}]|_{\mathbf{y}^{(k,R)} = \tilde{\mathbf{y}}^{(k,R)}} - y_{R,k} \mathbb{E}[x_{I,k} | \mathbf{y}]|_{\mathbf{y}^{(k,I)} = \tilde{\mathbf{y}}^{(k,I)}}. \quad (25)$$

The average $I^{(2)}(0) = \mathbb{E}_{\tilde{\mathbf{y}}} [I^{(2)}(0; \tilde{\mathbf{y}})]$ is obtained as

$$I^{(2)}(0) = \frac{4}{\sigma^4} \mathbb{E}_{\tilde{\mathbf{y}}} [\mathbf{d}^{(2)}(\tilde{\mathbf{y}}) \cdot \mathbf{1}]. \quad (26)$$

VI. COMPARISON OF CRB AND ACRB EXPRESSIONS

Here, we compare the ACRB expressions from (23) and (24)-(26) derived in the previous section to the exact CRB expression from (10)-(12).

A. Comparison of CRB and $\text{ACRB}^{(1)}$

Comparing (23) and (11), it follows that the first ad hoc computation method provides the correct CRB only when $\mathbf{D} = \mathbb{E}_{\mathbf{y}} [\mathbf{a}^T(\mathbf{y}) \mathbf{a}(\mathbf{y})]$ is a diagonal matrix. Hence, only if the sum of the diagonal elements in \mathbf{D} is large as compared to the sum of the off-diagonal elements, we have $I(0) \approx I^{(1)}(0)$ and $\text{ACRB}^{(1)}$ yields a close approximation of the true CRB.

We point out that if $x_{I,k} = 0$ for all k (e.g., pulse amplitude modulation (PAM) constellation), then $a_k(\mathbf{y}) = y_{I,k} \mathbb{E}[x_{R,k} | \mathbf{y}_R]$, $p(\mathbf{y}) = p(\mathbf{y}_R) p(\mathbf{y}_I)$ and $\mathbb{E}_{\mathbf{y}_I} [y_{I,k} y_{I,l}] = 0$ for $k \neq l$, so that automatically $D_{k,l} = 0$ for all $k \neq l$, yielding $\text{CRB} = \text{ACRB}^{(1)}$. Because $I(\theta)$ is independent of θ , the same holds if there exists an angle ψ such that $\text{Im}(x_k e^{j\psi}) = 0$ for any k (e.g., phase rotated PAM constellation).

Moreover, if the following holds for all $k \neq l$:

$$p(x_k, x_l | \mathbf{y}) = p(x_k | \mathbf{y}) p(x_l | \mathbf{y}), \quad (27)$$

then, for all $k \neq l$, $a_k(\mathbf{y}) a_l(\mathbf{y})$ equals the expectation of $\text{Im}(y_k x_k^*) \text{Im}(y_l x_l^*)$ with respect to $p(x_k, x_l | \mathbf{y})$. Consequently, $D_{k,l}$ from (12) reduces to $D_{k,l} = \mathbb{E}_{x_k, x_l} [E_k(x_k) E_l(x_l)]$ where $E_i(x_i) = \mathbb{E}_{y_i | x_i} [\text{Im}(y_i x_i^*)]$ for $i \in \{k, l\}$. Taking into account that $\mathbb{E}_{y_i | x_i} [y_i] = x_i$, it follows that $E_i(x_i) = 0$, yielding $D_{k,l} = 0$. Hence, (27) is a sufficient condition to have $I(0) = I^{(1)}(0)$, so that $\text{ACRB}^{(1)}$ equals the exact CRB expression.

B. Comparison of $\text{ACRB}^{(1)}$ and $\text{ACRB}^{(2)}$

Note that (22) and (26) in general yield different results. Indeed, (i) $a_k^{(2)}(y_k, \tilde{\mathbf{y}})$ from (25) contains $(\tilde{y}_{R,k}, \tilde{y}_{I,k})$, as opposed to $a_k^{(1)}(y_k, \tilde{\mathbf{y}}^{(k)})$ from (21); and (ii) the expectation in (24) is over the marginal conditional pdfs of $y_{R,k}$ and

$y_{I,k}$, as opposed to (20) where the expectation is over a joint conditional pdf of $y_{R,k}$ and $y_{I,k}$.

We point out in Appendix A that the following condition is sufficient to achieve $I^{(1)}(0) = I^{(2)}(0)$:

$$p(x_{R,k}, x_{I,k} | \mathbf{y}) = p(x_{R,k} | \mathbf{y}) p(x_{I,k} | \mathbf{y}). \quad (28)$$

Hence, when (28) holds, we have $\text{ACRB}^{(1)} = \text{ACRB}^{(2)}$. The condition (28) is trivially fulfilled, when the real and imaginary parts of \mathbf{x} are independently encoded, such that $p(\mathbf{x}) = p(\mathbf{x}_R) p(\mathbf{x}_I)$.

The sufficient conditions yielding $\text{ACRB}^{(1)} = \text{CRB}$ and $\text{ACRB}^{(1)} = \text{ACRB}^{(2)}$ are summarized in Table I below. Obviously, $\text{ACRB}^{(2)} = \text{CRB}$, in all cases where the condition for $\text{ACRB}^{(1)} = \text{ACRB}^{(2)}$ (second column) is simultaneously fulfilled with one of the conditions for $\text{ACRB}^{(1)} = \text{CRB}$ (first column).

$\text{ACRB}^{(1)} = \text{CRB}$	$\text{ACRB}^{(1)} = \text{ACRB}^{(2)}$
$p(x_k, x_l \mathbf{y})$ $= p(x_k \mathbf{y}) p(x_l \mathbf{y})$ or $\exists \psi \in [-\pi, \pi) : \text{Im}(\mathbf{x} e^{j\psi}) = \mathbf{0}$	$p(x_{R,k}, x_{I,k} \mathbf{y})$ $= p(x_{R,k} \mathbf{y}) p(x_{I,k} \mathbf{y})$

Table I
SUFFICIENT CONDITIONS FOR $\text{ACRB}^{(1)} = \text{CRB}$ AND FOR
 $\text{ACRB}^{(1)} = \text{ACRB}^{(2)}$.

VII. COMPUTATION OF CRB AND ACRBs

Let us now discuss the numerical evaluation of the above expressions for the CRB and ACRBs.

A. Conditional probabilities

The evaluation of the exact CRB, $\text{ACRB}^{(1)}$ and $\text{ACRB}^{(2)}$ requires the a posteriori pmfs $p(x_{R,k} | \mathbf{y})$ and $p(x_{I,k} | \mathbf{y})$ for $k = 1, \dots, K$ in the computation of $a_k(\mathbf{y})$ in (10) or $a_k^{(2)}(y_k, \tilde{\mathbf{y}})$ in (25).

The evaluation of $\text{ACRB}^{(2)}$ also requires the conditional pdfs $p(y_{R,k} | \mathbf{y}^{(k,R)})$ and $p(y_{I,k} | \mathbf{y}^{(k,I)})$ in (24). However, making use of

$$p(x_{R,k} | \mathbf{y}) \propto p(y_{R,k} | x_{R,k}) p(x_{R,k} | \mathbf{y}^{(k,R)}), \quad (29)$$

$$p(y_{R,k} | \mathbf{y}^{(k,R)}) = \sum_{x_{R,k}} p(y_{R,k} | x_{R,k}) p(x_{R,k} | \mathbf{y}^{(k,R)}), \quad (30)$$

where $p(x_{R,k} | \mathbf{y}^{(k,R)})$ is the extrinsic pmf of $x_{R,k}$, it follows immediately that the conditional pdfs $p(y_{R,k} | \mathbf{y}^{(k,R)})$ can be computed straightforwardly from the a posteriori pmfs $p(x_{R,k} | \mathbf{y})$. Similar expressions for $p(x_{I,k} | \mathbf{y})$ and $p(y_{I,k} | \mathbf{y}^{(k,I)})$ are obtained by replacing in (29) and (30) the label R by the label I.

In principle, the a posteriori pmfs $p(x_{R,k} | \mathbf{y})$ and $p(x_{I,k} | \mathbf{y})$ themselves can be obtained from a marginalization of $p(\mathbf{x} | \mathbf{y})$, with the latter computed from $p(\mathbf{y} | \mathbf{x})$ and $p(\mathbf{x})$ using Bayes' rule. However, in general the associated computational complexity increases exponentially with K . For many advanced

coded modulation schemes, a more computationally efficient approach (see e.g., [22]) consists of applying message passing on a factor graph (FG) representing a factorization of $p(\mathbf{x}_R, \mathbf{x}_I, \mathbf{e} | \mathbf{y})$, where $\mathbf{e} = (e_1, e_2, \dots)$ represents a set of additional variables (typically including the coded bits and/or the information bits), suitably chosen such that $p(\mathbf{x}_R, \mathbf{x}_I, \mathbf{e} | \mathbf{y})$ contains many factors, each containing a small number of variables. Using the sum-product algorithm (SPA) to generate the FG messages, an efficient marginalization of $p(\mathbf{x}_R, \mathbf{x}_I, \mathbf{e} | \mathbf{y})$ results (yielding, among others, the required $p(x_{R,k} | \mathbf{y})$ and $p(x_{I,k} | \mathbf{y})$ for $k = 1, \dots, K$), provided that the FG is a tree, i.e., the FG does not contain cycles. When the FG contains cycles, only an approximation, denoted as $p_{\text{FG}}(x_k | \mathbf{y})$, of $p(x_k | \mathbf{y})$ can be obtained via the SPA; the approximation is generally considered accurate when the cycles are long [23]. The main principles of the FG and SPA framework are summarized in Appendix B.

The computation in (24) involves a two-dimensional integration over $y_{R,k}$ and $y_{I,k}$. However, in Appendix C it is been pointed out (see also [18]) that by expanding the square in (24), three terms result, with each term being a product of a function depending on $y_{R,k}$ and a function depending on $y_{I,k}$. Integration of each individual term then results in a product of two one-dimensional integrals per term, which is computationally more efficient than a two-dimensional integral. On top of that, four out of six of these one-dimensional integrals can be evaluated analytically; the remaining two integrals can easily be computed numerically by using a quadrature rule.

B. Averaging over the observation

The correct and ad hoc expressions of the CRB can be expressed as $(\frac{4}{\sigma^4} \mathbb{E}_{\mathbf{y}} [Z^{(i)}(\mathbf{y})])^{-1}$, with $Z^{(0)}(\mathbf{y}) = \mathbf{1}^T \cdot \mathbf{a}^T(\mathbf{y}) \cdot \mathbf{a}(\mathbf{y}) \cdot \mathbf{1}$, $Z^{(1)}(\mathbf{y}) = \mathbf{a}(\mathbf{y}) \cdot \mathbf{a}^T(\mathbf{y})$ and $Z^{(2)}(\mathbf{y}) = \mathbf{d}^{(2)}(\mathbf{y}) \cdot \mathbf{1}$ yielding the correct CRB, $\text{ACRB}^{(1)}$ and $\text{ACRB}^{(2)}$, respectively.

The expectation $\mathbb{E}_{\mathbf{y}} [Z^{(i)}(\mathbf{y})]$ represents a $2K$ -dimensional integral over all K real and K imaginary components of \mathbf{y} . An approximation of $\mathbb{E}_{\mathbf{y}} [Z^{(i)}(\mathbf{y})]$ is obtained using Monte-Carlo integration. This involves (i) the independent generation of N_i coded symbol sequences $\{\mathbf{x}^{[1]}, \mathbf{x}^{[2]}, \dots, \mathbf{x}^{[N_i]}\}$ and the corresponding observation vectors $\{\mathbf{y}^{[1]}, \mathbf{y}^{[2]}, \dots, \mathbf{y}^{[N_i]}\}$ for $\theta = 0$, according to (1); (ii) the evaluation of $\{Z^{(i)}(\mathbf{y}^{[1]}), Z^{(i)}(\mathbf{y}^{[2]}), \dots, Z^{(i)}(\mathbf{y}^{[N_i]})\}$; and (iii) approximating $\mathbb{E}_{\mathbf{y}} [Z^{(i)}(\mathbf{y})]$ by the arithmetical average $Z_{\text{avg}}^{(i)} = \frac{1}{N_i} \sum_{n=1}^{N_i} Z^{(i)}(\mathbf{y}^{[n]})$. Denoting by $\text{MSE}^{(i)}$ the resulting MSE between $Z_{\text{avg}}^{(i)}$ and $\mathbb{E}_{\mathbf{y}} [Z^{(i)}(\mathbf{y})]$, we obtain $\text{MSE}^{(i)} = \frac{\text{VAR}^{(i)}}{N_i}$, where $\text{VAR}^{(i)}$ represents the variance of $Z^{(i)}(\mathbf{y})$; hence, the value of N_i that is required to achieve a given $\text{MSE}^{(i)}$ increases proportionally with $\text{VAR}^{(i)}$.

VIII. NUMERICAL RESULTS AND DISCUSSION

In this section we present numerical results regarding the exact and ad hoc CRB expressions. In all examples, unit-energy constellations are considered.

A. Coded modulation schemes

The following coded modulation schemes are considered:

(i) **BLO**: Three schemes use (6,2) linear block coding with code book $\mathbb{C} = \{(0,0,0,0,0,0), (0,0,1,1,1,1), (1,1,0,0,1,1), (1,1,1,1,0,0)\}$ yielding a minimum Hamming distance of 4, followed by 4-QAM Gray mapping. First, $\frac{2K}{3}$ information bits are encoded into the coded bit sequence $\mathbf{c} = (c_1, c_2, \dots, c_{2K})$, with $(c_{6n-5}, c_{6n-4}, \dots, c_{6n})$ belonging to \mathbb{C} for $n = 1, 2, \dots, \frac{K}{3}$. Subsequently, the bits from \mathbf{c} are Gray-mapped to 4-QAM symbols, yielding the symbol sequence $\mathbf{x} = (x_1, x_2, \dots, x_K)$, consisting of K symbols. We investigate three schemes, which differ in the way the symbol sequence \mathbf{x} is obtained from the coded bit sequence \mathbf{c} .

- In the BLO-1 scheme, the consecutive bits (c_{2k-1}, c_{2k}) are mapped to x_k , for $k = 1, \dots, K$. As the code book \mathbb{C} indicates that $c_{2k-1} = c_{2k}$, only two antipodal points from the 4-QAM constellation are effectively in use.
- In the BLO-2 scheme, the j th bit from the $(2l-1)$ th codeword and from the $2l$ th codeword are mapped to the real and imaginary part, respectively, of the symbol x_{6l-6+j} , for $j = 1, \dots, 6$ and $l = 1, \dots, K/6$. This way, $x_{R,k}$ and $x_{I,k}$ are statistically independent when conditioned on the observation \mathbf{y} .
- For the BLO-3 scheme, the coded bit sequence \mathbf{c} is applied to a pseudo-random interleaver of size $2K$, and the two consecutive bits at positions $2k-1$ and $2k$ at the interleaver output are mapped to the symbol x_k , for $k = 1, \dots, K$.

(ii) **REP**: A number of Km/n information bits is divided into K/n blocks of m information bits each; each block represent a M -ary information symbol, with $M = 2^m$. The information symbols are encoded using a M -ary $(n, 1)$ repetition code. Each of the resulting K coded M -ary symbols is Gray-mapped to a M -QAM symbol. Because of the $(n, 1)$ repetition code, the M -QAM symbol sequence consists of K/n strings of n identical symbols each, with symbols belonging to different strings being statistically independent and equiprobable.

(iii) **TUR**: The turbo coding and bit-interleaved coded modulation (BICM) schemes employed in [18]. The encoder is composed of two identical rate-1/2 recursive systematic convolutional (RSC) encoders that are concatenated in parallel via a pseudo-random information bit interleaver. Each constituent encoder is characterized by the generating polynomials $(1011, 1101)_2$ in binary notation, and a shift register that is in state 000 at the start of the encoding process; at the convolutional encoder input a tail bit sequence of length 3 resets the shift register to 000 at the end of the encoding process, and produces $2 \times 3 = 6$ tail bits at the convolutional encoder output. The systematic output of the second encoder is punctured, except for the 3 tail bits; hence, the turbo encoder output contains 12 tail bits in total. The scheme is applied to a sequence of $N_b = \frac{K}{3} \log_2 M - 4$ information bits and has a code rate of $r_c = \frac{N_b}{3N_b+12} (\approx 0.33$ for large N_b). The turbo encoder output is applied to a

pseudo-random coded bit interleaver and subsequently each $\log_2 M$ consecutive coded bits are Gray mapped to an M -ary square QAM (M -QAM) symbol. The scheme considered in [18] corresponds to $K = 207$.

(iv) **CON**: A sequence of $N_b = \frac{K}{2} \log_2 M - 3$ information bits is encoded using one of the rate-1/2 RSC codes constituting the above turbo code. Taking into account the 6 tail bits at the encoder output, this yields a code rate of $r_c = \frac{N_b}{2 \cdot N_b + 6} (\approx 0.5$ for large N_b). The resulting $K \log_2 (M)$ coded bits are Gray-mapped to K symbols from a rotated M -ary Phase Shift Keying (M -PSK) constellation. The constellation points are given by $\left\{ e^{j(\phi + k \frac{2\pi}{M})} : k = 0, 1, \dots, M-1 \right\}$ with ϕ denoting the rotation angle of the standard constellation. The rotated constellation is denoted M -PSK(ϕ); note that 4-PSK($\pi/4$) is identical to 4-QAM.

- In the CON-II scheme, a pseudo-random interleaver of size $K \log_2 M$ permutes the coded bits, prior to mapping $\log_2 M$ consecutive interleaver output bits to an M -PSK(ϕ) symbol.
- For the specific case of 4-QAM, we also consider an alternative CON scheme without interleaving, denoted CON-0, which directly maps each two consecutive encoder output bits to a data symbol.

B. Computation of a posteriori probabilities

For the BLO and REP coding schemes, the a posteriori pmfs $p(x_{R,k} | \mathbf{y})$ and $p(x_{I,k} | \mathbf{y})$ required to compute (10), (24) and (25) are straightforwardly obtained from the marginalization of $p(\mathbf{x}_R, \mathbf{x}_I | \mathbf{y})$. As $p(\mathbf{x}_R, \mathbf{x}_I)$ decomposes into the product of $\frac{K}{3}$ (BLO) or $\frac{K}{n}$ (REP) factors each containing only 6 (BLO) or $2n$ (REP) variables, with each element of $(\mathbf{x}_R, \mathbf{x}_I)$ occurring as a variable in exactly one factor, the computational complexity associated with this marginalization increases only linearly (rather than exponentially) with K .

For CON-0 (which uses 4-QAM), CON-II (for arbitrary PSK) and TUR (for arbitrary square QAM), the a posteriori pmfs $p(x_{R,k} | \mathbf{y})$ and $p(x_{I,k} | \mathbf{y})$ are replaced by the quantities $p_{\text{FG}}(x_{R,k} | \mathbf{y})$ and $p_{\text{FG}}(x_{I,k} | \mathbf{y})$ that result from applying the SPA to the FGs of Fig. 1(a)-(c), respectively. Here, $g_{R,k}(x_{R,k})$ and $g_{I,k}(x_{I,k})$ are short-hand notations of $p(y_{R,k} | x_{R,k})$ and $p(y_{I,k} | x_{I,k})$, respectively; the variables s_i and \hat{s}_i denote the state of the first and second RSC encoder after the i th trellis section; the variables b_m and c_n denote the m th information bit and the n th coded bit, respectively. The function $h_i(\cdot, \cdot, \dots)$ equals 1 if its variables satisfy the state transition constraints of the RSC encoder for the i th trellis section, and 0 otherwise; the functions $Q_k(\cdot, \cdot, \dots)$, $Q_{R,k}(\cdot, \cdot, \dots)$ and $Q_{I,k}(\cdot, \cdot, \dots)$ equal 1 when their variables satisfy the mapping constraints for the k th PSK symbol, for the real part of the k th QAM symbol and for the imaginary part of the k th QAM symbol, respectively, and 0 otherwise. It is worth noting that only for the special case of Gray-mapped 4-QAM the function $Q_k(\cdot, \cdot, \dots)$ can be further decomposed as:

$$Q_k(c_{k,1}, c_{k,2}, x_{R,k}, x_{I,k}) = Q_{R,k}(c_{k,1}, x_{R,k}) Q_{I,k}(c_{k,2}, x_{I,k}), \quad (31)$$

in which case in the FG from Fig. 1(b) the corresponding function node decomposes into two distinct function nodes as in Fig. 1(c).

The interleavers π , π_{inf} and π_{cod} , represented by the dotted rectangles, determine which variables appear in each factor, i.e., how the different function nodes are connected to the variable nodes. For both CON schemes, the SPA yields the exact a posteriori pmfs, i.e., $p_{\text{FG}}(x_{R,k}|\mathbf{y}) = p(x_{R,k}|\mathbf{y})$ and $p_{\text{FG}}(x_{I,k}|\mathbf{y}) = p(x_{I,k}|\mathbf{y})$, because the corresponding FGs contain no cycles. Note that, in the case of CON-0, the real and imaginary components of a given 4-QAM data symbol are related to the same trellis section, while, due to the interleaver π in Fig. 1(b), this is typically not the case with the CON-II scheme. As the FG for the TUR coding scheme contains cycles, the quantities $p_{\text{FG}}(x_{R,k}|\mathbf{y})$ and $p_{\text{FG}}(x_{I,k}|\mathbf{y})$ are only approximations of $p(x_{R,k}|\mathbf{y})$ and $p(x_{I,k}|\mathbf{y})$. However, the long interleavers in TUR ensure that the FG of Fig. 1(c) exhibits only long cycles, in which case $p_{\text{FG}}(x_{R,k}|\mathbf{y}) \approx p(x_{R,k}|\mathbf{y})$ and $p_{\text{FG}}(x_{I,k}|\mathbf{y}) \approx p(x_{I,k}|\mathbf{y})$.

Let us introduce the symbols x_k as additional variables in the FGs of Fig. 1 by adding factor nodes that represent the functions $q_k(x_k, x_{R,k}, x_{I,k})$, which equal 1 if $x_k = x_{R,k} + jx_{I,k}$ and 0 otherwise. The resulting augmented graphs, further denoted as FG_* , are shown in Fig. 2, where the additional nodes are indicated by means of a gray fill. Applying the SPA to any of these FG_* graphs, immediately yields among others $p_{\text{FG}_*}(x_{R,k}|\mathbf{y})$, $p_{\text{FG}_*}(x_{I,k}|\mathbf{y})$ and $p_{\text{FG}_*}(x_k|\mathbf{y})$, with

$$p_{\text{FG}_*}(x_k|\mathbf{y}) = p_{\text{FG}_*}(x_{R,k}|\mathbf{y})p_{\text{FG}_*}(x_{I,k}|\mathbf{y}), \quad (32)$$

for all k . Because all three FG_* graphs in Fig. 2 contain cycles, $p_{\text{FG}_*}(x_{R,k}|\mathbf{y})$, $p_{\text{FG}_*}(x_{I,k}|\mathbf{y})$ and $p_{\text{FG}_*}(x_k|\mathbf{y})$ are only approximations of the corresponding true a posteriori pmfs. However, the coded bit interleaver π_{cod} in Fig. 2(c) ensures that there are in general no short cycles in the FG_* graphs for TUR. Hence, for TUR, it is reasonable to assume that (32) implies that condition (28) holds approximately. A similar reasoning applies to the FG for 4-QAM CON-II (i.e., with $(M, \phi) = (4, \frac{\pi}{4})$) that results from decomposing the function nodes Q_k in Fig. 2(b) according to (31). This contrasts with Fig. 2(a) and Fig. 2(b), for which the cycles are very short; therefore, for CON-0 and for CON-II with $(M, \phi) \neq (4, \frac{\pi}{4})$, (32) does not imply (28).

We point out that, in a similar way, a factor graph $\text{FG}_{**}(k,l)$ can be constructed from the original FG, by merging for every k the variable nodes $x_{R,k}$, $x_{I,k}$ into a single node $x_k = (x_{R,k}, x_{I,k})$, and by introducing, for a particular (k, l) , the variable node $u_{k,l}$ and a function node representing the function $qq_k(u_{k,l}, x_k, x_l)$, which equals 1 if $u_{k,l} = (x_k, x_l)$ and 0 otherwise. The graphs $\text{FG}_{**}(k,l)$ corresponding to CON-0, to CON-II (for arbitrary PSK) and to TUR (for arbitrary square QAM) are shown in Fig. 3, where the (k, l) -dependent part of the graphs is indicated by means of double lines. Application of the SPA on these $\text{FG}_{**}(k,l)$ graphs yields $p_{\text{FG}_{**}(k,l)}(u_{k,l}|\mathbf{y}) = p_{\text{FG}_{**}(k,l)}(x_k|\mathbf{y})p_{\text{FG}_{**}(k,l)}(x_l|\mathbf{y})$. In all three cases, the cycles in the graphs $\text{FG}_{**}(k,l)$ are long, for the large majority of pairs (k, l) , so that $p(u_{k,l}|\mathbf{y}) \approx p(x_k|\mathbf{y})p(x_l|\mathbf{y})$, which indicates that condition (27) holds approximately.

C. Comparison of CRB and ACRBs

The independent generation of $N = 10^6$ observation vectors was performed to compute $I^{(i)}(0) = \frac{4}{\sigma^4} Z_{\text{avg}}^{(i)}$ and $\text{VAR}^{(i)}$ for $i = 0, 1, 2$ (see Section VII-B).

Figs. 4-7 show MCRB, CRB, $\text{ACRB}^{(1)}$ and $\text{ACRB}^{(2)}$ as a function of σ^{-2} in dB. More specifically,

- Fig. 4 depicts results for the BLO-1, BLO-2, BLO-3, CON-0, CON-II and TUR schemes, all with 4-QAM and $K = 207$.
- Fig. 5 is dedicated to REP, showing results for $M \in \{16, 64\}$, $n \in \{2, 20\}$ and $K = 200$. Here, $\text{MCRB}(\sigma^2)$, $\text{CRB}(\sigma^2; n)$ and $\text{ACRB}^{(i)}(\sigma^2; n)$ indicate the MCRB, the CRB and the ACRBs for given σ^2 and n .
- Results for CON-II with $K=100$, $M \in \{4, 8\}$ and $\phi \in \{0, \frac{\pi}{M}\}$ are displayed in Fig. 6.
- Fig. 7 corresponds to the TUR coding scheme with $K = 207$ and $M \in \{4, 16, 64, 256\}$.

To avoid overloaded figures, sets of (virtually) coinciding curves have been represented by only a single curve and legend entry; Table II indicates which of the depicted curves corresponds to CRB, $\text{ACRB}^{(1)}$ and $\text{ACRB}^{(2)}$ for the different coded modulation schemes.

Identifying from Table II the coded modulation schemes which achieve $\text{CRB} \approx \text{ACRB}^{(1)}$ and/or $\text{ACRB}^{(1)} \approx \text{ACRB}^{(2)}$, we obtain Table III. The results from Table III can be justified as follows:

- The results for TUR, CON-II, and CON-0 are consistent with the factor graph considerations from Section VIII-B, which indicate whether or not conditions (27) and/or (28) approximately hold. The result $\text{CRB} \approx \text{ACRB}^{(2)}$ for TUR agrees with the results presented in [18].
- As BLO-3 uses a pseudo-random coded-bit interleaver with a size much larger than the codeword length (and much larger than the number of bits per symbol) it is highly probable that $x_{R,k}$, $x_{I,k}$, $x_{R,l}$ and $x_{I,l}$ relate to different code words. Hence, (27) and (28) are likely to hold for the large majority of symbol indices, which explains why $\text{CRB} \approx \text{ACRB}^{(1)}$ and $\text{ACRB}^{(1)} \approx \text{ACRB}^{(2)}$.
- For BLO1, the only legitimate data symbols correspond to the bit patterns (0,0) and (1,1); using Gray mapping, the only allowed constellation points are antipodal (say, the points $e^{j\pi/4}$ and $-e^{j\pi/4}$). Hence, there exists an angle $\psi = -\pi/4$ yielding $\text{Im}(x_k e^{j\psi}) = 0$ for all k ; according to Remark 1, this is sufficient for $\text{CRB} = \text{ACRB}^{(1)}$ to hold. As only the constellation points $e^{j\pi/4}$ and $-e^{j\pi/4}$ are used, it is easily verified that condition (28) is not fulfilled for BLO-1; the corresponding $\text{ACRB}^{(1)}$ and $\text{ACRB}^{(2)}$ indeed yield substantially different curves.
- A significant deviation of $\text{ACRB}^{(1)}$ from the true CRB is observed only for coded modulation schemes involving symbol repetitions, as is the case for REP and for BLO-2; in the latter case, the 12 bits from the two codewords with indices $2l-1$ and $2l$ are mapped to the 4-QAM symbols $(x_{6l-5}, \dots, x_{6l})$, with $x_{6l-5} = x_{6l-4}$, $x_{6l-3} = x_{6l-2}$, $x_{6l-1} = x_{6l}$. If x_k is a priori known to equal x_l , then obviously $\mathbb{E}[x_k|\mathbf{y}] = \mathbb{E}[x_l|\mathbf{y}]$; hence, a high correlation between $a_k(\mathbf{y})$ and $a_l(\mathbf{y})$ can be expected, yielding

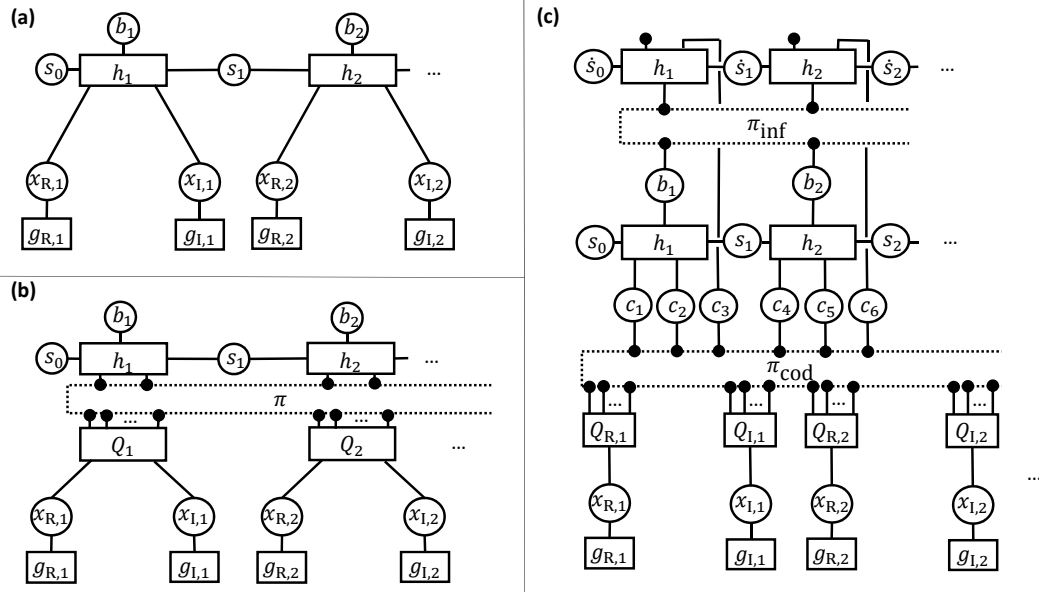


Figure 1. FG for the computation of the a posteriori pmfs $p(x_{R,k} | \mathbf{y})$ and $p(x_{I,k} | \mathbf{y})$; (a): CON-0, (b): CON-II and (c): TUR.

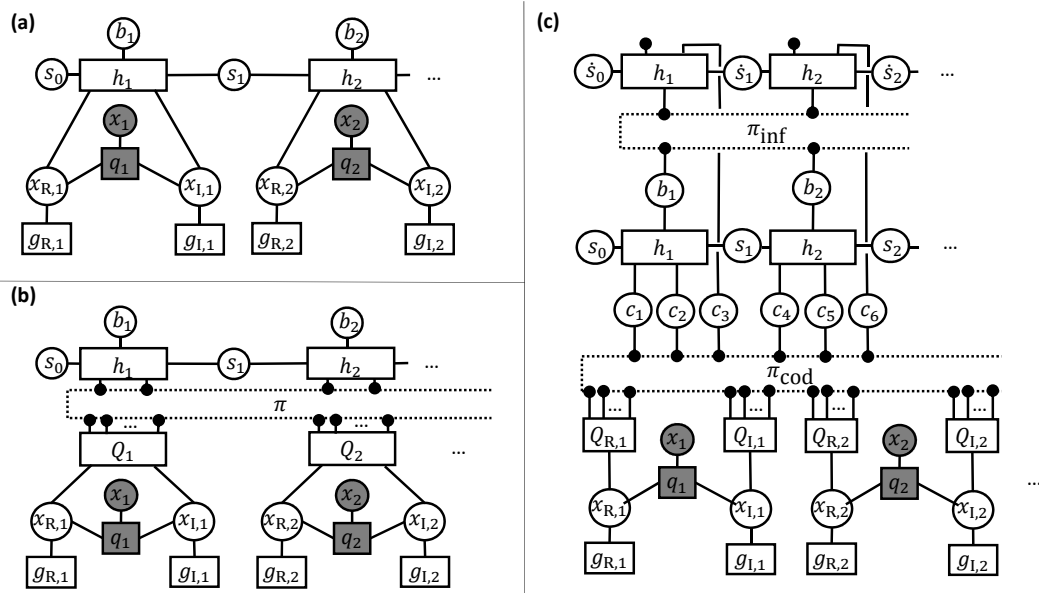


Figure 2. FG_* employed to check the validity of condition (28); (a): CON-0, (b): CON-II and (c): TUR.

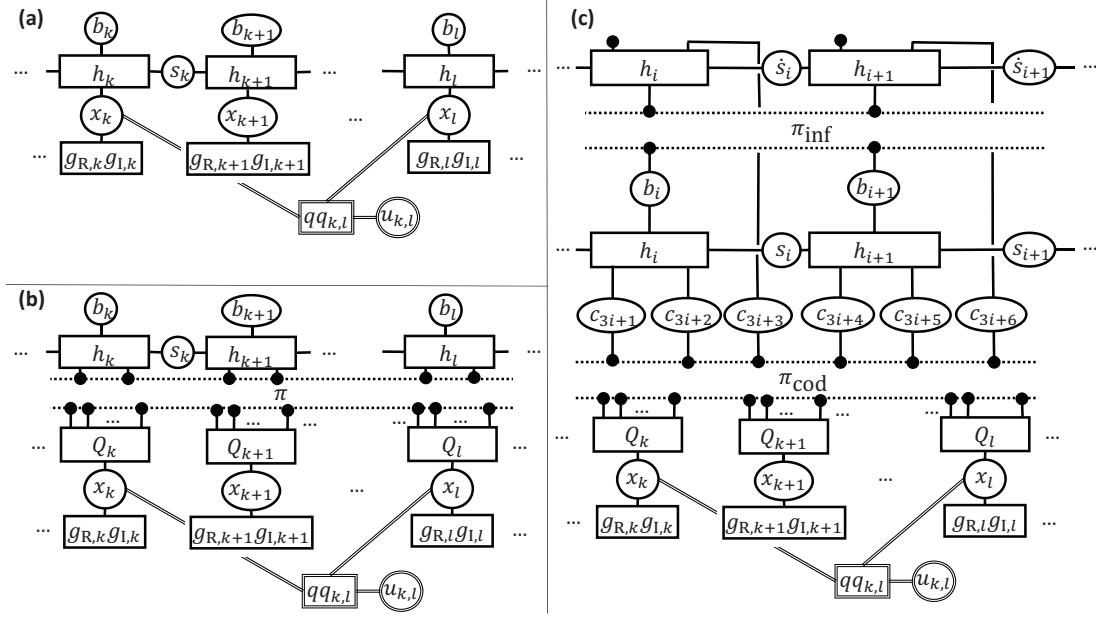


Figure 3. $FG_{**}(k,l)$ employed to check the validity of condition (27); (a): CON-0, (b): CON-II and (c): TUR.

Coded modulation	CRB	ACRB ⁽¹⁾	ACRB ⁽²⁾
BLO-1	BLO-1 CRB	BLO-1 CRB	BLO-3 CRB
BLO-2	BLO-2 CRB	BLO-3 CRB	BLO-3 CRB
BLO-3	BLO-3 CRB	BLO-3 CRB	BLO-3 CRB
REP	REP (M) CRB ($\sigma^2; n$)	REP (M) ACRB ($\sigma^2; n$)	REP (M) ACRB ($\sigma^2; n$)
TUR	TUR (M) CRB	TUR (M) CRB	TUR (M) CRB
CON-II, (M, ϕ) $\neq (4, \pi/4)$	CON-II (M) CRB	CON-II (M) CRB	CON-II (M, ϕ) ACRB ⁽²⁾
CON-II, (M, ϕ) = ($4, \pi/4$)	CON-II (4) CRB	CON-II (4) CRB	CON-II (4) CRB
CON-0	CON-0 CRB	CON-0 CRB	CON-II (4) CRB

Table II

LEGEND ENTRIES FOR FIG. 4, CORRESPONDING TO CRB, ACRB⁽¹⁾ AND ACRB⁽²⁾ FOR THE CONSIDERED CODED MODULATION SCHEMES.

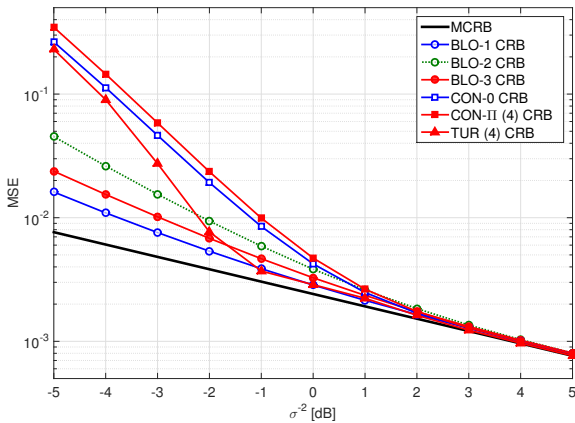


Figure 4. MCRB and (A)CRBs as a function of σ^{-2} , for 4-QAM and $K = 207$.

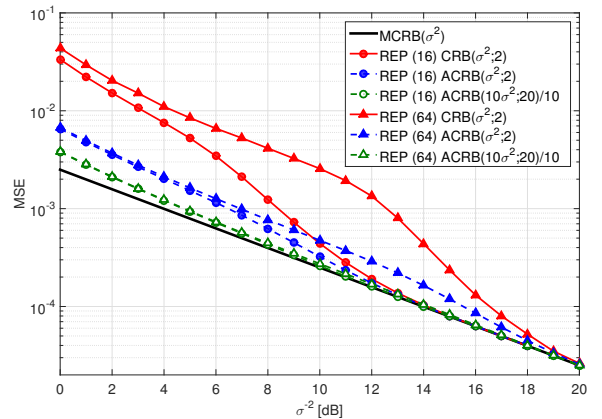


Figure 5. MCRB and (A)CRBs as a function of σ^{-2} , for REP and $K = 200$.

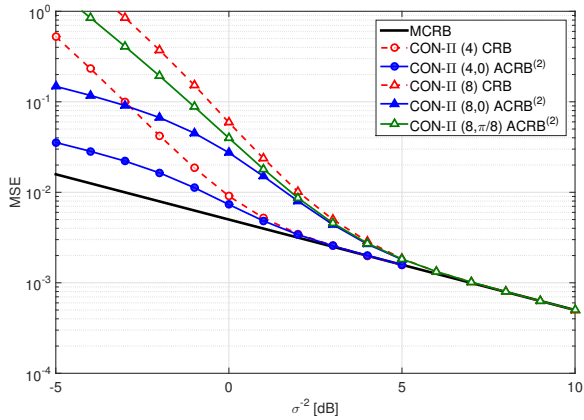


Figure 6. MCRB and (A)CRBs as a function of σ^{-2} , for CON-II and $K = 100$.

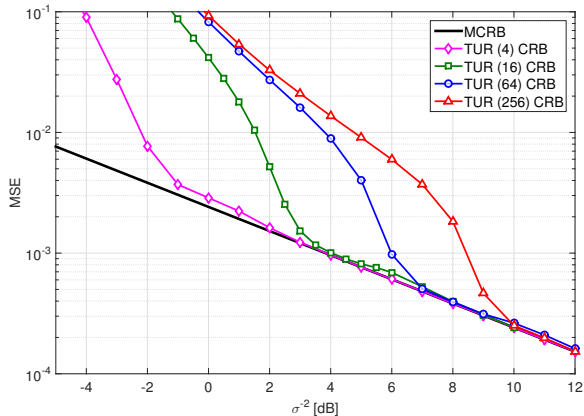


Figure 7. MCRB and (A)CRBs as a function of σ^{-2} , for TUR and $K = 207$.

$D_{k,l} \neq 0$ and therefore $\text{CRB} \neq \text{ACRB}^{(1)}$. It follows from Fig. 5 that the gap between $\text{ACRB}^{(1)}$ and CRB in the case of REP increases as the number n of identical symbols and/or the constellation size M increase.

- For BLO-2, the real and imaginary parts of any x_k correspond to different code words, such that condition (28) holds; this yields $\text{ACRB}^{(1)} = \text{ACRB}^{(2)}$.
- It is easily verified from the construction of REP that condition (28) holds; this yields $\text{ACRB}^{(1)} = \text{ACRB}^{(2)}$.

Coded modulation	$\text{CRB} \approx \text{ACRB}^{(1)}$	$\text{ACRB}^{(1)} \approx \text{ACRB}^{(2)}$
BLO-1	yes	no
BLO-2	no	yes
BLO-3	yes	yes
REP	no	yes
TUR	yes	yes
CON-II, $(M, \phi) \neq (4, \pi/4)$	yes	no
CON-II, $(M, \phi) = (4, \pi/4)$	yes	yes
CON-0	yes	no

Table III

COMPARISON OF CODED MODULATION SCHEMES.

Additional observations from Figs. 4-7 and Table II are the following:

- At high σ^{-2} , the CRB, $\text{ACRB}^{(1)}$ and $\text{ACRB}^{(2)}$ converge to the MCRB.
- In the case of 4-QAM modulation, all BLO schemes yield the same $\text{ACRB}^{(2)}$ expression. This follows immediately from (40)-(41) in Appendix C (Remark 3) and from the fact that the coded bit interleaver in front of the 4-QAM Gray mapper has no influence on the set of a posteriori pmfs values $\{p(x_{R,k} | \mathbf{y}), p(x_{I,k} | \mathbf{y})\}$ that is computed; the only effect of the interleaving operation is that the values in this set appear in a different order.³The same reasoning applies to all CON schemes that use 4-QAM.
- It can be verified that CRB related to the M -PSK(ϕ) constellation does not depend on the rotation angle ϕ . In contrast, the ACRBs for M -PSK(ϕ) do depend on ϕ .
- For REP, it is easily verified that $\text{CRB}(\sigma^2, n) = n\text{CRB}(\sigma^2/n, 1)$; hence, the curves for $\text{CRB}(\sigma^2, 2)$ and $\text{CRB}(10\sigma^2, 20)/10$ in Fig. 5 coincide (the legend entry only mentions $\text{CRB}(\sigma^2, 2)$). However, this property does not carry over to the ACRB. Indeed, Fig. 5 indicates that $\text{ACRB}(\sigma^2, n) \neq n\text{ACRB}(\sigma^2/n, 1)$ for $n > 1$, because the curves for $\text{ACRB}(\sigma^2, 2)$ and $\text{ACRB}(10\sigma^2, 20)/10$ corresponding to a given constellation are substantially different.

The above results show that for BLO-2 and REP, neither $\text{ACRB}^{(1)}$ nor $\text{ACRB}^{(2)}$ is a good approximation of CRB. For the other coded modulations we have considered, $\text{ACRB}^{(1)}$ is essentially the same as CRB. However, among those coded modulations yielding $\text{CRB} \approx \text{ACRB}^{(1)}$, only for TUR, for CON-II with 4-QAM and for BLO-3 we also obtain that $\text{ACRB}^{(2)}$ is essentially the same as CRB; these schemes have an interleaver between the encoder and the mapper, and independent mapping on the real and imaginary parts of the constellation. For the other coded modulations yielding $\text{CRB} \approx \text{ACRB}^{(1)}$ (i.e., CON-0, CON-II with $(M, \phi) \neq (4, \pi/4)$, BLO-1), $\text{ACRB}^{(2)}$ is substantially different from CRB. This indicates that $\text{ACRB}^{(1)}$ and $\text{ACRB}^{(2)}$ are close approximations of CRB only in specific cases, typically involving BICM combined with a binary code described by a factor graph without short cycles.

D. Maximum-likelihood estimator performance

In Fig. 8, the true and ad hoc CRBs for the BLO-1, BLO-2 and BLO-3 schemes with $K = 207$ are compared to the MSE performance of a state-of-the-art phase estimator that uses the expectation maximization (EM) algorithm to find the maximum likelihood (ML) estimate [13]. In our simulations, the true carrier phase θ was set equal to 0, and the estimate $\hat{\theta}(\mathbf{y}) = \hat{\theta}^{(100)}$ was obtained after 100 iterations, according to

$$\hat{\theta}^{(i)} = \arg \left(\sum_k y_k \sum_{m=0}^{M-1} \alpha_m^* p(x_k = \alpha_m | \mathbf{y}; \theta = \hat{\theta}^{(i-1)}) \right),$$

³This is because the value of the real and imaginary parts of a Gray-mapped 4-QAM symbol correspond to just a single coded bit.

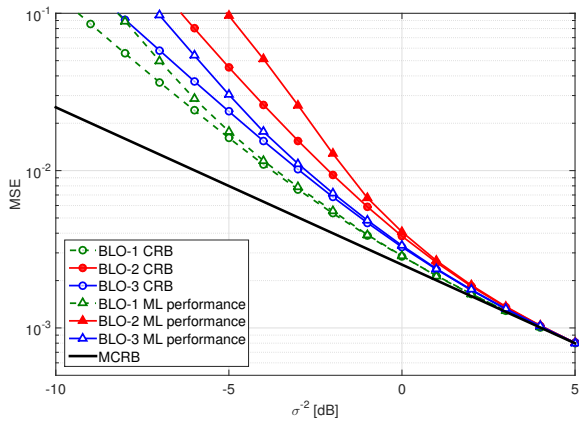


Figure 8. MCRB, (A)CRBs and MSE of EM-ML estimate as a function of σ^{-2} , for 4-QAM and $K = 207$.

for $i = 1, 2, \dots, 100$, with

$$\alpha_m = e^{j(\frac{\pi}{4} + m\frac{\pi}{2})}$$

and

$$\hat{\theta}^{(0)} = \frac{1}{4} \arg \left(- \sum_k (y_k)^4 \right),$$

where $\arg(a)$ denotes the phase angle of a complex number a , in the interval $[-\pi, \pi)$.

We make the following observations:

- The MSE resulting from the ML-based estimator is lower bounded by the true CRB in the considered range of MSE, for all three block coding schemes (BLO-1, BLO-2, BLO-3). This illustrates the usefulness of the CRB for phase estimation in a practical operating range involving small MSE.
- It follows from Table II that both ACRBs for BLO-2 are essentially the same and virtually coincide with the CRB for BLO-3, which is significantly lower than the CRB for BLO-2. Therefore, the MSE performance of the ML estimator cannot achieve the ACRBs.
- For BLO-1, Table II shows that $\text{ACRB}^{(1)}$ essentially coincides with the corresponding CRB, whereas $\text{ACRB}^{(2)}$ is virtually the same as the CRB for BLO-3, which is significantly higher than the CRB for BLO-1. In that case, the MSE performance of the ML estimator is significantly below $\text{ACRB}^{(2)}$.

These results illustrate that the true CRB provides a useful lower bound on the MSE resulting from practical estimators, whereas, depending on the coded modulation scheme, $\text{ACRB}^{(1)}$ and/or $\text{ACRB}^{(2)}$ can be close to the CRB, or significantly overestimate/underestimate the CRB.

E. Computational complexity

Only if $\text{CRB} \approx \text{ACRB}^{(i)}$ with $i \in \{1, 2\}$, it makes sense to quantify the computational advantage of evaluating $\text{ACRB}^{(i)}$ over the exact CRB, claimed in [18].

For the practically relevant case of advanced coded modulation schemes designed for FG-based detection, the time

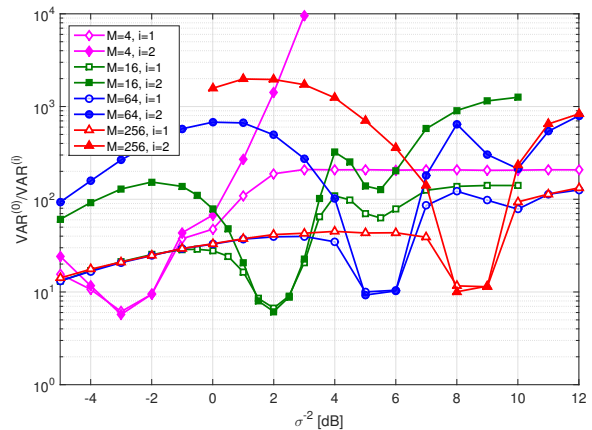


Figure 9. $\text{VAR}^{(0)}/\text{VAR}^{(i)}$ as a function of σ^{-2} for TUR and $K = 207$.

required to evaluate the quantity $Z_i(\mathbf{y})$ from Section VII-B (with $i = 0, 1, 2$ for the correct CRB, $\text{ACRB}^{(1)}$ and $\text{ACRB}^{(2)}$, respectively) for a given value of \mathbf{y} , is typically dominated by the time required to compute the a posteriori pmfs $p(x_{R,k}|\mathbf{y})$ and $p(x_{I,k}|\mathbf{y})$ for $k = 1, \dots, K$, which is independent of i . Hence, it follows from Section VII-B that the total time required to obtain the correct and ad hoc CRBs for a given value of $\text{MSE}^{(i)}$ is proportional to $\text{VAR}^{(i)}$.

Fig. 9 shows the ratio $\text{VAR}^{(0)}/\text{VAR}^{(i)}$ with $i = 1, 2$, for the TUR coding scheme considered in [18], for several values of the constellation size M ; we know from Table II that $\text{CRB} \approx \text{ACRB}^{(1)} \approx \text{ACRB}^{(2)}$ for this scheme. Considering that the computational complexities associated with the numerical evaluation of CRB, $\text{ACRB}^{(1)}$ and $\text{ACRB}^{(2)}$ for a given accuracy are proportional to $\text{VAR}^{(0)}$, $\text{VAR}^{(1)}$ and $\text{VAR}^{(2)}$, respectively, we conclude that

- For a given (σ^2, M) , either ACRB is simpler to evaluate than the exact CRB, and $\text{ACRB}^{(2)}$ is simpler to evaluate than $\text{ACRB}^{(1)}$.
- For given (i, M) , the smallest difference in complexity between $\text{ACRB}^{(i)}$ and CRB is observed for σ^2 corresponding to $\text{CRB} \approx 3 \cdot \text{MCRB}$, while the largest gain is observed for values of σ^2 at which the CRB approaches the MCRB. We note that the latter is easily evaluated exactly as $\sigma^2/(2K) = \sigma^2/414$.
- The computational advantage of $\text{ACRB}^{(i)}$ over the exact CRB varies significantly as a function of σ^2 , i and M . For $M \in \{4, 16, 64, 256\}$, Table IV lists the minimum and maximum values of $\text{VAR}^{(0)}/\text{VAR}^{(i)}$ (as identified from Fig. 9) for $i \in \{1, 2\}$, within the range of interest for σ^{-2} . This range of interest is the interval of σ^{-2} for which the CRB is smaller than 10^{-1} and strictly larger than the MCRB. (As this range is derived from Fig. 7, values of σ^{-2} larger than 12 dB are not considered.) The smallest and largest values of $\text{VAR}^{(0)}/\text{VAR}^{(i)}$ observed are 6 (for $M = 4$, $i = 1$ or 2, and $\sigma^{-2} = -3$ dB) and 10^4 (for $M = 4$, $i = 2$ and $\sigma^{-2} = 3$ dB), respectively.

M	range of σ^{-2}	VAR ⁽⁰⁾ /VAR ⁽¹⁾		VAR ⁽⁰⁾ /VAR ⁽²⁾	
		min.	max.	min.	max.
4	[-4,3] (dB)	6	200	6	10000
16	[-1,8] (dB)	7	120	7	900
64	[0,12] (dB)	9	120	9	700
256	[0,12] (dB)	10	120	10	700

Table IV
COMPUTATIONAL ADVANTAGE OF EVALUATING ACRB⁽ⁱ⁾, WITH $i = 1, 2$,
OVER THE EXACT CRB.

IX. CONCLUSIONS AND REMARKS

In this paper, we presented a detailed comparison of the analytical expressions for the exact CRB and for two ad hoc CRBs, in the context of non-Bayesian carrier phase estimation from a channel-coded, linearly modulated signal corrupted by AWGN. We derived sufficient conditions (regarding the a posteriori probability of the data symbols) which guarantee that the expressions for the CRB and the ad hoc CRBs are identical. We explained that these sufficient conditions are essentially met for BICM systems using codes with long cycles and QAM constellations with independent in-phase and quadrature mapping; hence, for these types of coded modulations the ad hoc CRBs are a valid alternative to the CRB.

The CRB and the ad hoc CRBs were numerically evaluated for various coded modulation systems. We pointed out that, for some of these coded modulations, the two ad hoc CRBs yield essentially the same numerical result as the CRB, whereas, for other coded modulations, one or both of the ad hoc CRBs yield numerical values which considerably differ from the CRB.

We also investigated the MSE of the ML phase estimator. For practical operating conditions, the CRB is observed to be a useful lower bound on this MSE, whereas, depending on the coded modulation considered, one or both of the ad hoc CRBs can be substantially larger or smaller than the MSE; hence, in general one must be careful when using either of the ad hoc CRBs as a substitute for the CRB, because there is no guarantee that the ad hoc CRB even lower bounds the MSE.

We compared the computational complexities related to the numerical evaluation of the CRB and the ad hoc CRBs, and pointed out the smaller complexity associated with the latter.

For ease of exposition, we focused on the simple case of non-Bayesian carrier phase estimation. However, these results can straightforwardly be extended to the non-Bayesian joint parameter (symbol timing, frequency offset, phase offset, signal amplitude, ...) estimation, with the random coded symbols as nuisance parameters. As future work, other (more complex) scenarios for parameter estimation from a channel-coded, linearly modulated signal corrupted by AWGN could be envisaged; these scenarios include the Bayesian (or mixed Bayesian and non-Bayesian) estimation in the presence of deterministic and/or random nuisance parameters (e.g., the joint estimation of the symbol timing and the frequency offset, where the nuisance parameters are the random coded data symbols and the signal amplitude and carrier phase), for which other types of CRBs (such as the Bayesian CRB [4] and the hybrid CRB [24], [25]) must be considered.

APPENDIX A

Let us assume that the sufficient condition (28) holds. Dividing both sides from (28) by the product of $p(y_{R,k} | x_{R,k})$ and $p(y_{I,k} | x_{I,k})$ yields

$$p(x_{R,k}, x_{I,k} | \mathbf{y}^{(k)}) = p(x_{R,k} | \mathbf{y}^{(k,R)}) p(x_{I,k} | \mathbf{y}^{(k,I)}).$$

Summing both sides over $x_{I,k}$, one obtains

$$p(x_{R,k} | \mathbf{y}^{(k)}) = p(x_{R,k} | \mathbf{y}^{(k,R)}). \quad (33)$$

As $p(x_{R,k} | \mathbf{y}) \propto p(y_{R,k} | x_{R,k}) p(x_{R,k} | \mathbf{y}^{(k,R)})$, it follows from (33) that $p(x_{R,k} | \mathbf{y})$ does not depend on $y_{I,k}$. Similarly, it can be shown that $p(x_{I,k} | \mathbf{y})$ does not depend on $y_{R,k}$. Hence, comparing (21) and (25), we have $a_k^{(2)}(y_k, \tilde{\mathbf{y}}) = a_k^{(1)}(y_k, \tilde{\mathbf{y}}^{(k)})$.

Taking into account that

$$p(x_{R,k}, y_{R,k} | \mathbf{y}^{(k,R)}) \propto p(y_{R,k} | x_{R,k}) p(x_{R,k} | \mathbf{y}^{(k,R)}),$$

it follows from (33) that $p(x_{R,k}, y_{R,k} | \mathbf{y}^{(k,R)})$ does not depend on $y_{I,k}$, implying that $p(y_{R,k} | \mathbf{y}^{(k,R)}) = p(y_{R,k} | \mathbf{y}^{(k)})$. As a result, we have

$$p(y_{R,k}, y_{I,k} | \mathbf{y}^{(k)}) = p(y_{R,k} | \mathbf{y}^{(k,R)}) p(y_{I,k} | \mathbf{y}^{(k,I)}),$$

which indicates that the averaging in (20) is the same as the averaging in (24).

APPENDIX B

Assuming that a joint conditional pmf $p(\mathbf{v} | \mathbf{y})$ can be factorized as $p(\mathbf{v} | \mathbf{y}) = \prod_i F_i(v_{k_{1,i}}, v_{k_{2,i}}, \dots, v_{k_{N_i,i}})$, the FG and SPA framework can be briefly summarized as follows [22]:

(i) In the FG, the function node F_i is connected to the variable nodes $v_{k_{1,i}}, v_{k_{2,i}}, \dots, v_{k_{L_i,i}}$ and the SPA computes the messages $\mathcal{M}_{F_i \rightarrow v_{k_{l,i}}}(v_{k_{l,i}})$ from F_i to $v_{k_{l,i}}$, for $l \in \{1, 2, \dots, L_i\}$, as:

$$\mathcal{M}_{F_i \rightarrow v_{k_{l,i}}}(v_{k_{l,i}}) \propto \sum_{v_{k_{j,i}}; j \neq l} F_i(v_{k_{1,i}}, v_{k_{2,i}}, \dots, v_{k_{N_i,i}}) \prod_{j \neq l} \mathcal{M}_{v_{k_{j,i}} \rightarrow F_i}(v_{k_{j,i}}). \quad (34)$$

(ii) Assume that a variable v_k occurs in the factors $F_{i_{j,k}}(\dots, v_k, \dots)$ for $j = 1, 2, \dots, n_k$. Then, in the FG, the variable node v_k is connected to the function nodes $F_{i_{1,k}}, F_{i_{2,k}}, \dots, F_{i_{n_k,k}}$. The SPA computes the messages $\mathcal{M}_{v_k \rightarrow F_{i_{l,k}}}(v_k)$ from v_k to $F_{i_{l,k}}$, for $l \in \{1, 2, \dots, n_l\}$, as:

$$\mathcal{M}_{v_k \rightarrow F_{i_{l,k}}}(v_k) \propto \begin{cases} \prod_{j \neq l} \mathcal{M}_{F_{i_{j,k}} \rightarrow v_k}(v_k) & , n_k > 1 \\ 1 & , n_k = 1 \end{cases}. \quad (35)$$

where $\mathcal{M}_{F_{i_{j,k}} \rightarrow v_k}(v_k)$ denotes the message from $F_{i_{j,k}}$ to v_k .

(iii) An approximation $p_{\text{FG}}(v_k | \mathbf{y})$ of the a posteriori pmf $p(v_k | \mathbf{y})$ of a FG variable v_k is computed as:

$$p_{\text{FG}}(v_k | \mathbf{y}) \propto \mathcal{M}_{F_{i_{j,k}} \rightarrow v_k}(v_k) \mathcal{M}_{v_k \rightarrow F_{i_{j,k}}}(v_k), \quad (36)$$

with $F_{i_{j,k}}$ any desired function node connected to v_k and $(\mathcal{M}_{F_{i_{j,k}} \rightarrow v_k}(v_k), \mathcal{M}_{v_k \rightarrow F_{i_{j,k}}}(v_k))$ the two messages (one in each direction) along the corresponding edge.

APPENDIX C

Expanding the square in (24), we obtain

$$\begin{aligned} d_k^{(2)}(\tilde{\mathbf{y}}) &= \alpha_{R,k} \left(\tilde{\mathbf{y}}^{(k,R)} \right) \mathbb{E}_{y_{I,k} | \tilde{\mathbf{y}}^{(k,I)}} \left[\beta_{I,k} \left(y_{I,k}, \tilde{\mathbf{y}}^{(k,I)} \right) \right] \\ &\quad - 2\gamma_{R,k} \left(\tilde{\mathbf{y}}^{(k,R)} \right) \gamma_{I,k} \left(\tilde{\mathbf{y}}^{(k,I)} \right) \\ &\quad + \alpha_{I,k} \left(\tilde{\mathbf{y}}^{(k,R)} \right) \mathbb{E}_{y_{R,k} | \tilde{\mathbf{y}}^{(k,R)}} \left[\beta_{R,k} \left(y_{R,k}, \tilde{\mathbf{y}}^{(k,R)} \right) \right], \end{aligned} \quad (37)$$

where

$$\begin{aligned} \alpha_{R,k} \left(\tilde{\mathbf{y}}^{(k,R)} \right) &= \mathbb{E}_{y_{R,k} | \tilde{\mathbf{y}}^{(k,R)}} \left[(y_{R,k})^2 \right] \\ &= \mathbb{E}_{x_{R,k} | \tilde{\mathbf{y}}^{(k,R)}} \left[\mathbb{E} \left[(y_{R,k})^2 | x_{R,k} \right] \right] \\ &= \mathbb{E}_{x_{R,k} | \tilde{\mathbf{y}}^{(k,R)}} \left[(x_{R,k})^2 \right] + \frac{\sigma^2}{2}, \end{aligned} \quad (38)$$

$\beta_{R,k} \left(y_{R,k}, \tilde{\mathbf{y}}^{(k,R)} \right)$ is a short hand notation for $\left(\mathbb{E} [x_{R,k} | \mathbf{y}] \Big|_{\mathbf{y}^{(k,R)} = \tilde{\mathbf{y}}^{(k,R)}} \right)^2$ and

$$\begin{aligned} \gamma_{R,k} \left(\tilde{\mathbf{y}}^{(k,R)} \right) &= \mathbb{E}_{y_{R,k}, x_{R,k} | \tilde{\mathbf{y}}^{(k,R)}} [y_{R,k} x_{R,k}] \\ &= \mathbb{E}_{x_{R,k} | \tilde{\mathbf{y}}^{(k,R)}} [x_{R,k} \mathbb{E} [y_{R,k} | x_{R,k}]] \\ &= \mathbb{E}_{x_{R,k} | \tilde{\mathbf{y}}^{(k,R)}} \left[(x_{R,k})^2 \right]. \end{aligned} \quad (39)$$

The expressions for $\alpha_{I,k} \left(\tilde{\mathbf{y}}^{(k,I)} \right)$, $\beta_{I,k} \left(y_{I,k}, \tilde{\mathbf{y}}^{(k,I)} \right)$ and $\gamma_{I,k} \left(\tilde{\mathbf{y}}^{(k,I)} \right)$ follow from replacing in (38)-(39) the label R by the label I. The computation of (37) involves only two one-dimensional integrations, over $y_{R,k}$ and $y_{I,k}$, respectively.

Remark 1. For the special case of a 4-QAM constellation, we have $(x_{R,l})^2 = (x_{R,k})^2 = (x_{I,k})^2 = \frac{1}{2}$ for all l and k , such that $\mathbb{E}_{x_{R,k} | \tilde{\mathbf{y}}^{(k,R)}} \left[(x_{R,k})^2 \right] = \mathbb{E}_{x_{I,k} | \tilde{\mathbf{y}}^{(k,I)}} \left[(x_{I,k})^2 \right] = \frac{1}{2}$ in (38) and (39). As a result, the average $I^{(2)}(0) = \mathbb{E}_{\tilde{\mathbf{y}}} [I^{(2)}(0; \tilde{\mathbf{y}})]$ in (26) simplifies to:

$$I^{(2)}(0) = \frac{4}{\sigma^4} \left(\bar{\mathbf{d}}^{(2)} \cdot \mathbf{1} \right), \quad 4\text{-QAM} \quad (40)$$

with $\bar{\mathbf{d}}^{(2)} = \left(\bar{d}_1^{(2)}, \bar{d}_2^{(2)}, \dots, \bar{d}_K^{(2)} \right)$ and

$$\begin{aligned} \bar{d}_k^{(2)} &= \frac{1 + \sigma^2}{2} \mathbb{E}_{y_{R,k}, \tilde{\mathbf{y}}^{(k,R)}} \left[\beta_{R,k} \left(y_{R,k}, \tilde{\mathbf{y}}^{(k,R)} \right) \right] - \frac{1}{2} \\ &\quad + \frac{1 + \sigma^2}{2} \mathbb{E}_{y_{I,k}, \tilde{\mathbf{y}}^{(k,I)}} \left[\beta_{I,k} \left(y_{I,k}, \tilde{\mathbf{y}}^{(k,I)} \right) \right]. \end{aligned} \quad (41)$$

ACKNOWLEDGMENT

This research has been funded by the Inter-university Attraction Poles Program initiated by the Belgian Science Policy Office. The first author would like to acknowledge the financial support of the Fund for Scientific Research - Flanders.

REFERENCES

- [1] N. Noels and M. Moeneclaey, "Synchronization of digital signals," in *Transmission Techniques for Digital Communications*, pp. 227–281, Elsevier BV, 2016.
- [2] H. Cramér, *Mathematical methods of statistics*. Princeton, 1946.
- [3] C. R. Rao, "Information and accuracy attainable in the estimation of statistical parameters," *Bull. Calcutta Math. Soc.*, vol. 37, no. 3, pp. 81–91, 1945.
- [4] H. L. Van Trees, *Detection, Estimation, and Modulation Theory-Part I-Detection, Estimation, and Linear Modulation Theory*. John Wiley & Sons New York, 2001.
- [5] A. N. D'Andrea, U. Mengali, and R. Reggiannini, "The modified Cramér-Rao bound and its application to synchronization problems," *IEEE Transactions on Communications*, vol. 42, pp. 1391–1399, Feb. 1994.
- [6] M. Moeneclaey, "On the true and the modified Cramér-Rao bounds for the estimation of a scalar parameter in the presence of nuisance parameters," *IEEE Transactions on Communications*, vol. 46, pp. 1536–1544, Nov. 1998.
- [7] F. Rice, B. Cowley, B. Moran, and M. Rice, "Cramér-Rao lower bounds for QAM phase and frequency estimation," *IEEE Transactions on Communications*, vol. 49, pp. 1582–1591, Sept. 2001.
- [8] N. Noels, H. Steendam, and M. Moeneclaey, "The true Cramér-Rao bound for carrier frequency estimation from a PSK signal," *IEEE Transactions on Communications*, vol. 52, pp. 834–844, May 2004.
- [9] J. P. Delmas, "Closed-form expressions of the exact Cramér-Rao bound for parameter estimation of BPSK, MSK, or QPSK waveforms," *IEEE Signal Processing Letters*, vol. 15, pp. 405–408, 2008.
- [10] F. Bellili, N. Atallah, S. Affes, and A. Stephenne, "Cramér-Rao lower bounds for frequency and phase NDA estimation from arbitrary square QAM-modulated signals," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4517–4525, 2010.
- [11] H. Steendam and M. Moeneclaey, "Low-SNR limit of the Cramér-Rao bound for estimating the carrier phase and frequency of a PAM, PSK, or QAM waveform," *IEEE Communications Letters*, vol. 5, pp. 218–220, May 2001.
- [12] N. Noels, H. Steendam, and M. Moeneclaey, "The Cramér-Rao bound for phase estimation from coded linearly modulated signals," *IEEE Communications Letters*, vol. 7, pp. 207–209, May 2003.
- [13] N. Noels, H. Steendam, and M. Moeneclaey, "Carrier and clock recovery in (turbo) coded systems: Cramér-Rao bound and synchronizer performance," *Appl. Signal Process., EURASIP J.*, vol. 2005, pp. 972–980, May 2005.
- [14] C. Herzet, N. Noels, V. Lottici, H. Wymeersch, M. Luise, M. Moeneclaey, and L. Vandendorpe, "Code-aided turbo synchronization," *Proceedings of the IEEE*, vol. 95, pp. 1255–1271, June 2007.
- [15] J. Yang, B. Geller, and A. Wei, "Approximate expressions for Cramér-Rao bounds of code aided QAM dynamical phase estimation," in *Communications, 2009. ICC'09. IEEE International Conference on*, pp. 1–5, IEEE, 2009.
- [16] N. Wu, H. Wang, H. Zhao, H. Liu, and J. Kuang, "Evaluation of Cramér-Rao bounds for phase estimation of coded linearly modulated signals," in *Proc. IEEE 79th Vehicular Technology Conf. (VTC Spring)*, pp. 1–5, May 2014.
- [17] F. Bellili, A. Methenni, and S. Affes, "Closed-form CRLBs for SNR estimation from turbo-coded BPSK-, MSK-, and square-QAM-modulated signals," *IEEE Transactions on Signal Processing*, vol. 62, no. 15, pp. 4018–4033, 2014.
- [18] F. Bellili, A. Methenni, and S. Affes, "Closed-form CRLBs for CFO and phase estimation from turbo-coded square-QAM-modulated transmissions," *IEEE Transactions on Wireless Communications*, vol. 14, pp. 2513–2531, May 2015.
- [19] F. Bellili, C. Elguet, S. B. Amor, S. Affes, and A. Stéphenne, "Code-aided DOA estimation from turbo-coded QAM transmissions: Analytical CRLBs and maximum likelihood estimator," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 2850–2865, 2017.
- [20] F. Bellili, A. Methenni, S. B. Amor, S. Affes, and A. Stephenne, "Code-aided time synchronization of turbo-coded square-QAM-modulated transmissions: Closed-form Cramér-Rao lower bounds," in *Global Communications Conference (GLOBECOM), 2015 IEEE*, pp. 1–7, IEEE, 2015.
- [21] T. Rountenberg and J. Tabrikian, "Non-bayesian periodic Cramér-Rao bound," *IEEE Transactions on Signal Processing*, vol. 61, no. 4, pp. 1019–1032, 2013.
- [22] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, 2001.
- [23] J. Wainwright, M., T. S. Jaakkola, and A. S. Willsky, "Tree-based reparameterization framework for analysis of sum-product and related algorithms," *Trans. Inf. Th., IEEE*, vol. 49, pp. 1120–1146, May 2003.
- [24] Y. Noam and H. Messer, "Notes on the tightness of the hybrid Cramér-Rao lower bound," *IEEE Transactions on Signal Processing*, vol. 57, no. 6, pp. 2074–2084, 2009.
- [25] S. Bar and J. Tabrikian, "Bayesian estimation in the presence of deterministic nuisance parameters-Part I: Performance bounds," *IEEE Transactions on Signal Processing*, vol. 63, no. 24, pp. 6632–6646, 2015.