

# Code-Aided Time Synchronization of Turbo-Coded Square-QAM-Modulated Transmissions: Closed-Form Cramér-Rao Lower Bounds

Fauzi Bellili, Achref Methenni, Souheib Ben Amor, Sofiène Affes, and Alex Stéphenne  
INRS-EMT, 800, de la Gauchetière Ouest, Bureau 6900, Montreal, Qc, H5A 1K6, Canada.  
Emails: {bellili, methenni, souheib.ben.amor, affes}@emt.inrs.ca, and stephenne@ieee.org

**Abstract**—This paper tackles the problem of code-aided (CA) timing recovery in turbo-coded square-QAM transmissions. Owing to a simple recursive construction process, some hidden properties of Gray-coded (GC) square-QAM constellations are demystified and the code bits’ *a priori* log-likelihood ratios (LLRs) are explicitly incorporated in the log-likelihood function (LLF). Then, by splitting the underlying LLF into the sum of two analogous terms, we derive for the very first time the closed-form expressions for the exact Cramér-Rao lower bounds (CRLBs) of the underlying turbo synchronization problem. Computer simulations will show that the new closed-form CRLBs coincide exactly with their empirical counterparts evaluated previously using exhaustive Monte-Carlo simulations. They will also show unambiguously the remarkable performance improvements of the CA scheme against the traditional non-data-aided (NDA) one. Over a wide range of practical SNRs, the new CA CRLBs reach those of the *completely* data-aided (DA) scheme in which all the transmitted symbols are perfectly known to the receiver.

## I. INTRODUCTION

**I**N order to provide high quality of service while satisfying the ever-increasing demand in high data rates, the use of powerful error-correcting codes in conjunction with high-spectral-efficiency modulations is advocated. Indeed, turbo codes along with high-order quadrature amplitude modulations (QAMs) are two key features of current and future wireless communication standards such as 4G long-term evolution (LTE), LTE-advanced (LTE-A) and beyond (LTE-B) [1, 2]. As a crucial task in any digital receiver [3], accurate time synchronization remains a challenging problem especially for turbo-coded systems since they are intended to operate at very low signal-to-noise ratios (SNRs). In fact, the widespread adoption of turbo codes is in part fueled by their ability to operate in the near-Shannon limit even under such adverse SNR conditions [4]. Yet, the salutary performance of these powerful error-correcting codes is prone to severe degradations if the system is not accurately synchronized in time, phase or frequency. The goal of time synchronization, in particular, consists in estimating and compensating for the unknown time delay (TD) introduced by the channel in an attempt to provide the decision device with signal samples of lowest possible ISI corruption [3].

Generally speaking, depending on the *a priori* knowledge about the transmitted symbols, time delay estimators (TDEs) can be categorized into two broad categories: DA and NDA methods assuming, respectively, all the symbols to be perfectly known or completely unknown (and thus considered as nuisance parameters) to the receiver. DA TDEs have the major drawback of limiting the whole throughput of the system as they require the regular transmission of a completely known (i.e., pilot) symbol sequence. NDA TDEs, however, suffer from severe performance degradations under harsh SNR conditions. It is essential then,

to strike a proper balance between these two extreme cases and CA estimation sprang up exactly for this purpose. In this paper, we consider the problem of CA time synchronization from a “performance bounds” point of view. We typically consider the *stochastic*<sup>1</sup> CRLB which, unlike many other loose bounds, reflects the actual achievable performance in practice. Yet, even under uncoded transmissions, the complex structure of the LLF makes it extremely hard, if not impossible, to derive analytical expressions for this practical bound, especially with high-order modulations. Therefore, in the specific context of timing recovery, the stochastic CRLBs were previously evaluated using exhaustive Monte-Carlo simulations (i.e., *empirically*) in [11] and [10] for both non-code-aided (NCA) and CA estimations, respectively. Just recently though were their analytical expressions established [12] but only in the NCA case.

Motivated by these facts, we develop in this paper for the very first time the closed-form expressions of CA TD CRLBs with arbitrary square-QAM modulations. Compared to their NCA counterparts [11, 12], the new closed-form CA CRLBs confirm that huge performance gains, in terms of TD estimation, can be indeed achieved by appropriately exploiting the decoder’s *soft* outputs during the estimation process. The closed-form TD CA CRLBs are obtained owing to an elegant decomposition of the system’s LLF into the sum of two analogous terms after exploiting some hidden properties of the the Gray coding process.

The rest of this paper is structured as follows. In section II, we present the system model. In section III, we introduce the recursive process that allows the construction of any GC square-QAM constellation and relate the symbols’ *a priori* probabilities (APPs) to the *a priori* LLRs of the code bits. In section IV, we derive the LLF of the system and in Section V we establish the new closed-form CA CRLBs. In Section VI, we provide and discuss some graphical representations of the obtained bounds. Finally, we draw out some concluding remarks in section VII.

We mention beforehand that some of the common notations will be used in this paper. Vectors and matrices are represented in lower- and upper-case bold fonts, respectively. The operators  $\Re\{\cdot\}$  and  $\Im\{\cdot\}$  return, respectively, the real and imaginary parts of any complex number. The operators  $\{\cdot\}^*$  and  $|\cdot|$  return its conjugate and its amplitude, respectively, and  $j$  is the pure complex number that verifies  $j^2 = -1$ . The Kronecker and Dirac delta functions are denoted, respectively, as  $\delta_{m,n}$  and  $\delta(t)$ . We will also denote the probability mass function (PMF) for discrete random variables (RVs) by  $P[\cdot]$  and the pdf for continuous RVs by  $p[\cdot]$ . The statistical expectation is denoted as  $\mathbb{E}\{\cdot\}$  and the notation  $\triangleq$  is used for definitions.

<sup>1</sup>In linearly-modulated transmissions, the term *stochastic* refers to estimation under the assumption of unknown and *random* transmitted symbols.

## II. SYSTEM MODEL

Consider a turbo-coded system where a binary sequence of information bits is fed into a turbo encoder consisting of two identical recursive and systematic convolutional codes (RSCs) which are concatenated in parallel via an inner interleaver. The resulting code bits are fed into a puncturer which selects an appropriate combination of the parity bits, from both encoders, in order to achieve a desired code rate  $R$ . The code bit sequence is then scrambled with an outer interleaver,  $\Pi_2$ , and mapped onto a given GC constellation. The obtained symbols,  $\{a(k)\}_{k=1}^K$ , are pulse-shaped and the corresponding continuous-time signal:

$$x(t) = \sum_{k=1}^K a(k) h(t - kT), \quad (1)$$

is transmitted over the communication channel where  $T$  and  $h(t)$  are the symbol duration and a unit-energy square-root shaping pulse, respectively. The transmitted symbols  $\{a(k)\}_k$  which are completely unknown to the receiver are drawn from any  $M$ -ary GC square-QAM constellation whose alphabet is denoted as  $\mathcal{C}_p = \{c_0, c_1, \dots, c_{M-1}\}$ . Here, by *square* QAM we mean  $M = 2^{2p}$  (i.e., QPSK, 16-QAM, 64-QAM, etc...). The Nyquist pulse  $g(t)$  obtained from  $h(t)$  is defined as:

$$g(t) = \int_{-\infty}^{+\infty} h(x)h(t+x)dx, \quad (2)$$

and satisfies the first Nyquist criterion [3]:

$$g(nT) = 0, \text{ for any integer } n \neq 0. \quad (3)$$

At the receiver side, assuming perfect frequency and phase synchronizations, the delayed continuous-time received signal before matched filtering is expressed as:

$$y(t) = \sqrt{E_s} x(t - \tau) + w(t), \quad (4)$$

where  $E_s$  is the transmit signal energy and  $\tau$  is the unknown time delay parameter to be estimated. Moreover,  $w(t)$  is a *proper* complex additive white Gaussian noise (AWGN) with independent real and imaginary parts, each of variance  $\sigma^2$  (i.e., with overall noise power  $N_0 = 2\sigma^2$ ). The SNR of the channel is also defined as:

$$\rho \triangleq \frac{E_s}{2\sigma^2}. \quad (5)$$

An integral step in the derivation of CRLBs consists in finding the LLF of the system. This requires an averaging of the unknown symbols over the constellation alphabet. In *completely* NDA estimation (or before data detection), no *a priori* information is available about the transmitted symbols. Therefore, the latter are usually assumed to be equally likely, i.e.,  $P[a(k) = c_m] = \frac{1}{M}$  for  $k = 1, 2, \dots, K, \forall c_m \in \mathcal{C}_p$ . In CA estimation, however, the actual APPs of the symbols must be used in order to enhance the estimation performance. Indeed, in the next section, we will express these APPs in terms of the *a priori* LLRs of the individual conveyed bits. As will be explained later in Section IV-A, accurate estimates for these LLRs (and in turn the symbols' APPs) can be obtained in practice from the soft outputs of the two SISO decoders at the convergence of the BCJR algorithm [18].

As mentioned earlier, consider a GC square-QAM constellation wherein each point,  $\{c_m\}_{m=1}^M$ , of the underlying alphabet is mapped onto a unique sequence of  $\log_2(M) = 2p$  bits denoted here as  $\bar{b}_1^m \bar{b}_2^m \dots \bar{b}_l^m \dots \bar{b}_{2p}^m$ . For the sake of clarity, this mapping will be denoted as follows

$$c_m \longleftrightarrow \bar{b}_1^m \bar{b}_2^m \dots \bar{b}_l^m \dots \bar{b}_{2p}^m. \quad (6)$$

The same notation is used to refer to the  $k^{\text{th}}$  bit sequence,  $b_1^k b_2^k \dots b_l^k \dots b_{2p}^k$ , that is conveyed by the transmission of the symbol  $x(k)$ , i.e.,  $x(k) \longleftrightarrow b_1^k b_2^k \dots b_l^k \dots b_{2p}^k$ . Due to the large-size interleaver, the code bits can be reasonably assumed as statistically independent. This assumption is indeed pervasive in CA estimation practices (see [13-17] and references therein). Therefore, the APPs of each transmitted symbol,  $x(k)$ , factorize into the elementary APPs of its conveyed bits:

$$P[x(k) = c_m] = \prod_{l=1}^{2p} P[b_l^k = \bar{b}_l^m], \quad 1 \leq m \leq M. \quad (7)$$

We also define the *a priori* LLR of the  $l^{\text{th}}$  code bit,  $b_l^k$ , conveyed by  $x(k)$  as follows:

$$L_l(k) \triangleq \ln \left( \frac{P[b_l^k = 1]}{P[b_l^k = 0]} \right). \quad (8)$$

Using (8) and the fact that  $P[b_l^k = 0] + P[b_l^k = 1] = 1$ , it can be easily shown that:

$$P[b_l^k = \bar{b}_l^m] = \frac{1}{2 \cosh(L_l(k)/2)} e^{(2\bar{b}_l^m - 1) \frac{L_l(k)}{2}}, \quad (9)$$

in which  $\bar{b}_l^m$  is either 0 or 1 depending on which of the symbols  $c_m$  is transmitted, at time instant  $k$ , and of course on the Gray mapping associated to the constellation in (6). Thus, injecting (9) in (7) the symbols' APPs develop into:

$$P[x(k) = c_m] = \underbrace{\prod_{l=1}^{2p} \frac{1}{2 \cosh(L_l(k)/2)}}_{\beta_k} \prod_{l=1}^{2p} e^{(2\bar{b}_l^m - 1) \frac{L_l(k)}{2}}. \quad (10)$$

Next, we describe a simple three-step process that allows the recursive construction of arbitrary GC square-QAM constellations. Owing to this recursive process, some hidden properties of such constellations will be revealed and carefully handled in order to rewrite the APPs in a more insightful form allowing for LF factorization in the next section. In fact, using any *basic* GC QPSK and starting from any given GC  $2^{2(p-1)}$ -QAM, it is possible to build another GC  $2^{2p}$ -QAM as follows:

- *step 1*: build the top-right quadrant of the desired  $2^{2p}$ -QAM from all the points of the available  $2^{2(p-1)}$ -QAM. As such, all the points of the new quadrant are still missing two out of the  $2p$  bits they must represent. This is simply because they were cloned from the given  $2^{2(p-1)}$ -QAM whose points actually represent  $2p - 2$  bits only. For the sake of clarity and again w.l.o.g, we will assume that these two missing bits always occupy the two least significant positions.
- *step 2*: build the remaining empty quadrants (i.e., top-left, bottom-right, and bottom-left) of the desired  $2^{2p}$ -QAM by symmetries with respect to the  $y$ -axis, the  $x$ -axis, and the center point, respectively. In light of "*step 1*", all the points of the desired constellation are inherently missing two bits each.
- *step 3*: copy the two bits of each quadrant in the basic GC QPSK constellation to all the points that belong to the same quadrant in the incomplete  $2^{2p}$ -QAM constellation obtained in "*step 2*".

One example that clearly depicts these three steps is shown in Fig. 1 (illustrated here in initial transition from 4-QAM to 16-QAM). For the sake of clarity and conciseness, however, and again w.l.o.g, we will from now on consider the GC QPSK depicted in Fig. 1) as the *basic* one and will actually use it in

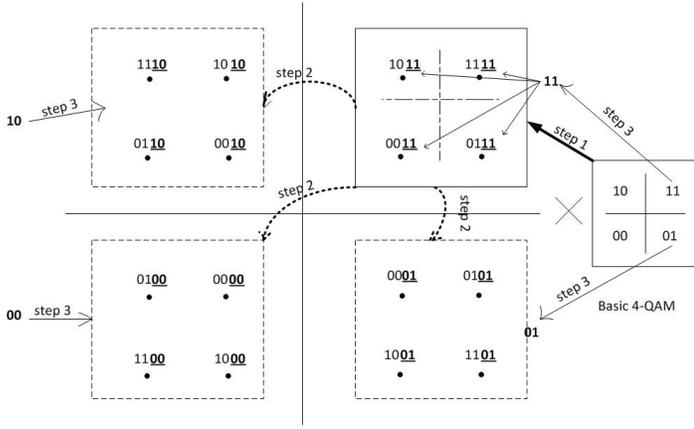


Fig. 1. Recursive construction of GC square-QAM constellations illustrated here from 4-QAM to 16-QAM.

both the *initial* and all subsequent construction iterations required for a given modulation order. For the same reasons, we will also assume that the two missing bits in “*step 1*” are always the two LSBs.

By recalling that  $\mathcal{C}_p$  denotes the whole alphabet of the obtained GC  $2^{2p}$ -QAM constellation, we further denote its top-right quadrant by  $\tilde{\mathcal{C}}_p$ . As such, each point  $c_m \in \mathcal{C}_p$  belongs to a set of four symmetrical points  $\{\tilde{c}_m, \tilde{c}_m^*, -\tilde{c}_m, -\tilde{c}_m^*\}$  for some  $\tilde{c}_m \in \tilde{\mathcal{C}}_p$ . Moreover, due to symmetries of “*step 2*”, these four symmetrical points share the same  $2p - 2$  most significant bits (MSBs),  $\bar{b}_1^m \bar{b}_2^m \bar{b}_3^m \dots \bar{b}_{2p-3}^m \bar{b}_{2p-2}^m$ . We use them to define the quantity:

$$\mu_{k,p}(c_m) \triangleq \prod_{l=1}^{2p-2} e^{(2\bar{b}_l^m - 1) \frac{L_l(k)}{2}}, \quad \forall c_m \in \mathcal{C}_p, \quad (11)$$

which verifies:

$$\mu_{k,p}(\tilde{c}_m) = \mu_{k,p}(-\tilde{c}_m) = \mu_{k,p}(\tilde{c}_m^*) = \mu_{k,p}(-\tilde{c}_m^*). \quad (12)$$

Using (11) in (10) it follows that for  $\forall c_m \in \mathcal{C}_p$  we have:

$$P[x(k) = c_m] = \beta_k \mu_{k,p}(c_m) e^{(2\bar{b}_{2p-1}^m - 1) \frac{L_{2p-1}(k)}{2}} e^{(2\bar{b}_{2p}^m - 1) \frac{L_{2p}(k)}{2}}. \quad (13)$$

Moreover, the two bits  $\bar{b}_{2p-1}^m$  and  $\bar{b}_{2p}^m$  involved in the two exponentials of (13) are exactly the same for all the symbols that belong to the same quadrant in the obtained  $2^{2p}$ -QAM GC constellation (recall that they are added in “*step 3*”). Typically, by considering the *basic* QPSK depicted in Fig. 1, they are given by “ $\bar{b}_{2p-1}^m \bar{b}_{2p}^m$ ” = “11”, “00”, “01” and “10”, for each  $\tilde{c}_m \in \tilde{\mathcal{C}}_p$ ,  $-\tilde{c}_m \in -\tilde{\mathcal{C}}_p$ ,  $\tilde{c}_m^* \in \tilde{\mathcal{C}}_p^*$  and  $-\tilde{c}_m^* \in -\tilde{\mathcal{C}}_p^*$ , respectively. Therefore, by injecting these results in (13) and using (12), it follows that  $\forall \tilde{c}_m \in \tilde{\mathcal{C}}_p$  we have:

$$Pr[x(k) = \tilde{c}_m] = \beta_k \mu_{k,p}(\tilde{c}_m) e^{\frac{L_{2p-1}(k)}{2}} e^{\frac{L_{2p}(k)}{2}}, \quad (14)$$

$$Pr[x(k) = \tilde{c}_m^*] = \beta_k \mu_{k,p}(\tilde{c}_m) e^{-\frac{L_{2p-1}(k)}{2}} e^{\frac{L_{2p}(k)}{2}}, \quad (15)$$

$$Pr[x(k) = -\tilde{c}_m] = \beta_k \mu_{k,p}(\tilde{c}_m) e^{-\frac{L_{2p-1}(k)}{2}} e^{-\frac{L_{2p}(k)}{2}}, \quad (16)$$

$$Pr[x(k) = -\tilde{c}_m^*] = \beta_k \mu_{k,p}(\tilde{c}_m) e^{\frac{L_{2p-1}(k)}{2}} e^{-\frac{L_{2p}(k)}{2}}. \quad (17)$$

### III. DERIVATION OF THE LLF

After some algebraic manipulations detailed in [21], it can be shown that the *conditional* (on the transmitted symbols) LLF of

the system is given by (see also [12] for a prior work on NDA TD CRLBs):

$$\Lambda(\mathbf{y}|\mathbf{a}; \tau) = \exp\left\{\frac{\sqrt{E_s}}{\sigma^2} \int_{\mathbb{R}} \Re\{y(t)x(t-\tau)^*\} dt - \frac{E_s}{2\sigma^2} \int_{\mathbb{R}} |x(t-\tau)|^2 dt\right\}, \quad (18)$$

where  $\mathbf{a} = [a(1), a(2), \dots, a(K)]^T$  and  $\mathbf{y}$  is an infinite-dimensional vector that contains the projection coefficients of  $y(t)$  over an orthonormal basis for the finite-energy signals subspace. Now, replacing the transmitted signal  $x(t)$  by its expression given in (1), and using the fact that the shaping pulse,  $h(t)$ , verifies the first-order Nyquist criterion, it can be shown that:

$$\Lambda(\mathbf{y}|\mathbf{a}; \tau) = \prod_{k=1}^K \Omega_\tau(a(k), y(t)). \quad (19)$$

where

$$\Omega_\tau(a(k), y(t)) \triangleq \exp\left\{\frac{\sqrt{E_s}}{\sigma^2} \int_{\mathbb{R}} \Re\{y(t)a(k)^*\} h(t-kT-\tau) dt - \frac{E_s}{2\sigma^2} |a(k)|^2\right\}. \quad (20)$$

The *unconditional* LF,  $\Lambda(\mathbf{y}; \tau)$ , is obtained by averaging (19) over all the possible transmitted symbol sequences of size  $K$ , i.e.,  $\Lambda(\mathbf{y}; \tau) = \mathbb{E}_{\mathbf{a}}\{\Lambda(\mathbf{y}|\mathbf{a}; \tau)\}$  leading to:

$$\Lambda(\mathbf{y}; \tau) = \sum_{\mathbf{c}_i \in \mathcal{C}_p^K} P[\mathbf{a} = \mathbf{c}_i] \Lambda(\mathbf{y}|\mathbf{a} = \mathbf{c}_i; \tau). \quad (21)$$

Since the coded bits are assumed to be statistically independent, the transmitted symbols are also independent thereby leading to:

$$P[\mathbf{a} = \mathbf{c}_i] = \prod_{k=1}^K P[a(k) = c_i(k)]. \quad (22)$$

Plugging (19) and (22) in (21), it can be shown that:

$$\begin{aligned} \Lambda(\mathbf{y}; \tau) &= \sum_{\mathbf{c}_i \in \mathcal{C}_p^K} \prod_{k=1}^K P[a(k) = c_i(k)] \Omega_\tau(c_i(k), y(t)) \\ &= \prod_{k=1}^K \sum_{c_m \in \mathcal{C}_p} P[a(k) = c_m] \Omega_\tau(c_m, y(t)). \end{aligned} \quad (23)$$

Therefore, the *unconditional* LLF defined as  $\mathcal{L}(\mathbf{y}; \tau) \triangleq \ln(\Lambda(\mathbf{y}; \tau))$ , is given by:

$$\mathcal{L}(\mathbf{y}; \tau) = \sum_{k=1}^K \ln(\bar{\Omega}_k(\tau, y(t))), \quad (24)$$

in which  $\bar{\Omega}_k(\tau, y(t))$  is simply the average of  $\Omega_\tau(a(k), y(t))$  over the constellation alphabet, i.e.:

$$\bar{\Omega}_k(\tau, y(t)) \triangleq \sum_{c_m \in \mathcal{C}_p} P[a(k) = c_m] \Omega_\tau(c_m, y(t)). \quad (25)$$

In the following, we will further manipulate the term  $\bar{\Omega}_k(\tau, y(t))$  and ultimately factorize it into the product of two analogous terms which involve two independent and *almost* identically distributed RVs. As a starting point, notice that the alphabet of any  $M$ -ary QAM constellation with  $M = 2^{2p}$  (i.e., square-QAM) is expressed in the  $I/Q$  plane as  $\mathcal{C}_p = \{\pm(2i-1)d_p \pm j(2n-1)d_p\}_{i,n=1,2,\dots,2^{p-1}}$ , where  $2d_p$  is the minimum inter-symbol distance<sup>2</sup>. Moreover, owing to the constellation

<sup>2</sup>For normalized-energy constellations, we have  $\frac{1}{2^{2p}} \sum_{m=1}^{2^{2p}} |c_m|^2 = 1$  from which the expression of  $d_p$  is obtained as  $d_p = 2^{p-1} / \sqrt{2^p \sum_{m=1}^{2^{2p-1}} (2m-1)^2}$ .

symmetry, we have  $\mathcal{C}_p = \tilde{\mathcal{C}}_p \cup (-\tilde{\mathcal{C}}_p) \cup \tilde{\mathcal{C}}_p^* \cup (-\tilde{\mathcal{C}}_p^*)$  where  $\tilde{\mathcal{C}}_p = \{(2i-1)d_p + j(2n-1)d_p\}_{i,n=1,2,\dots,2^{p-1}}$  is its top-right quadrant. Thus, noticing that  $|\tilde{c}_m| = |-\tilde{c}_m| = |\tilde{c}_m^*| = |-\tilde{c}_m^*|$ , and recalling (20), it follows that (25) is equivalent to:

$$\begin{aligned} \bar{\Omega}_k(\tau, y(t)) &= \sum_{\tilde{c}_m \in \tilde{\mathcal{C}}_p} e^{-\frac{E_s}{2\sigma^2} |\tilde{c}_m|^2} \times \\ &\left( P[a(k) = \tilde{c}_m] \exp \left\{ \frac{\sqrt{E_s}}{\sigma^2} \int_{\mathbb{R}} \Re\{\tilde{c}_m^* y(t)\} h(t-kT-\tau) dt \right\} \right. \\ &+ P[a(k) = -\tilde{c}_m] \exp \left\{ \frac{\sqrt{E_s}}{\sigma^2} \int_{\mathbb{R}} \Re\{-\tilde{c}_m^* y(t)\} h(t-kT-\tau) dt \right\} \\ &+ P[a(k) = \tilde{c}_m^*] \exp \left\{ \frac{\sqrt{E_s}}{\sigma^2} \int_{\mathbb{R}} \Re\{\tilde{c}_m y(t)\} h(t-kT-\tau) dt \right\} \\ &\left. + P[a(k) = -\tilde{c}_m^*] \exp \left\{ \frac{\sqrt{E_s}}{\sigma^2} \int_{\mathbb{R}} \Re\{-\tilde{c}_m y(t)\} h(t-kT-\tau) dt \right\} \right). \end{aligned}$$

For ease of notations, we will no longer show the dependence of  $\bar{\Omega}_k(\tau, y(t))$  on the received signal,  $y(t)$ , and denote it simply as  $\bar{\Omega}_k(\tau)$ . Therefore, after using the explicit expressions of the four symbols' APPs already established in (14) to (17) and the trivial identity  $e^x + e^{-x} = 2 \cosh(x)$ , it can be shown that:

$$\begin{aligned} \bar{\Omega}_k(\tau) &= 2\beta_k \sum_{\tilde{c}_m \in \tilde{\mathcal{C}}_p} \mu_{k,p}(\tilde{c}_m) e^{-\rho |\tilde{c}_m|^2} \times \\ &\left[ \cosh \left\{ \frac{\sqrt{E_s}}{\sigma^2} \int_{\mathbb{R}} \Re\{\tilde{c}_m y(t)\} h(t-kT-\tau) dt + \frac{L_{2p}(k) - L_{2p-1}(k)}{2} \right\} + \right. \\ &\left. \cosh \left\{ \frac{\sqrt{E_s}}{\sigma^2} \int_{\mathbb{R}} \Re\{\tilde{c}_m^* y(t)\} h(t-kT-\tau) dt + \frac{L_{2p}(k) + L_{2p-1}(k)}{2} \right\} \right]. \end{aligned}$$

Furthermore, by using the relationship  $\cosh(x) + \cosh(y) = 2 \cosh\left(\frac{x+y}{2}\right) \cosh\left(\frac{x-y}{2}\right)$  along with the two identities  $\tilde{c}_m + \tilde{c}_m^* = 2\Re\{\tilde{c}_m\}$  and  $\tilde{c}_m - \tilde{c}_m^* = 2j\Im\{\tilde{c}_m\}$ , it can be shown that:

$$\begin{aligned} \bar{\Omega}_k(\tau) &= 4\beta_k \sum_{\tilde{c}_m \in \tilde{\mathcal{C}}_p} \mu_{k,p}(\tilde{c}_m) e^{-\rho |\tilde{c}_m|^2} \cosh \left\{ \frac{\sqrt{E_s} \Re\{\tilde{c}_m\}}{\sigma^2} u_k(\tau) + \frac{L_{2p}(k)}{2} \right\} \\ &\times \cosh \left\{ \frac{\sqrt{E_s} \Im\{\tilde{c}_m\}}{\sigma^2} v_k(\tau) + \frac{L_{2p-1}(k)}{2} \right\}, \quad (26) \end{aligned}$$

in which  $u_k(\tau)$  and  $v_k(\tau)$  are the matched-filtered *in-phase* and *quadrature* components of the received signal given by:

$$u_k(\tau) = \int_{-\infty}^{+\infty} \Re\{y(t)\} h(t-kT-\tau) dt, \quad (27)$$

$$v_k(\tau) = \int_{-\infty}^{+\infty} \Im\{y(t)\} h(t-kT-\tau) dt. \quad (28)$$

Since each  $\tilde{c}_m \in \tilde{\mathcal{C}}_p$  can be written as  $\tilde{c}_m = [2i-1]d_p + j[2n-1]d_p$  for some  $1 \leq i, n \leq 2^{p-1}$ , then the single sum over  $\tilde{c}_m$  in (26) can be equivalently replaced by a double sum over the two counters  $i$  and  $n$ . Therefore, if we are able to factorize  $\mu_{k,p}(\tilde{c}_m) = \mu_{k,p}([2i-1]d_p + j[2n-1]d_p)$  as the product of two terms, one depending on  $i$  only and the other on  $n$  only, then  $\bar{\Omega}_k(\tau)$  will be factorized as well by splitting the two sums (over  $i$  and  $n$ ). To see this, we will rather use<sup>3</sup> the superscript  $(i, n)$  instead of  $m$  for the bits,  $\bar{b}_l^m$ , that are associated to  $c_m$ , i.e.:

$$\bar{b}_l^m \equiv \bar{b}_l^{(i,n)}, \quad \text{for } l = 1, 2, \dots, 2p. \quad (29)$$

<sup>3</sup>Normally, we should use  $(i_m, n_m)$  but, for ease of notation, we drop the index  $m$  and simply use  $(i, n)$ .

We also show by mathematical induction the following extremely useful result for the required decomposition of  $\mu_{k,p}(\tilde{c}_m)$ .

**LEMMA 1:** The obtained  $2^{2p}$ -QAM GC constellation has the following property:

- The odd-position bits,  $\bar{b}_{2l-1}^{(i,n)}$ , do not change by scanning each *horizontal* line of the constellation points.
- The even-position bits,  $\bar{b}_{2l}^{(i,n)}$ , do not change by scanning each *vertical* line of the constellation points.

In a nutshell, the fact that the odd-position bits,  $\bar{b}_{2l-1}^{(i,n)}$ , do not change for each horizontal line means that they do not change by varying the symbols' abscissa,  $(2i-1)d_p$ , or equivalently by changing the counter  $i$ . Therefore,  $\{\bar{b}_{2l-1}^{(i,n)}\}_{l=1}^p$  are function of  $n$  only and, by the same token, the even-position bits,  $\{\bar{b}_{2l}^{(i,n)}\}_{l=1}^p$ , are function of  $i$  only. As a consequence, we will from now on drop the vanishing counter from each group of bits, i.e.:

$$\left\{ \bar{b}_{2l-1}^{(i,n)} \equiv \bar{b}_{2l-1}^{(n)} \quad \text{and} \quad \bar{b}_{2l}^{(i,n)} \equiv \bar{b}_{2l}^{(i)} \right\}_{l=1}^p. \quad (30)$$

By exploiting the result above in (11), i.e.:

$$\mu_{k,p}(\tilde{c}_m) \triangleq \prod_{l=1}^{2p-2} e^{(2\bar{b}_l^{(i,n)} - 1) \frac{L_l(k)}{2}}, \quad \forall \tilde{c}_m \in \tilde{\mathcal{C}}_p, \quad (31)$$

and then rearranging the odd-position (resp. even-position) bits together, we split  $\mu_{k,p}(\tilde{c}_m)$  into the product of two separate terms each of which depends solely on one of the two counters  $i$  and  $n$ :

$$\mu_{k,p}(\tilde{c}_m) = \theta_{k,2p}^{(i)} \theta_{k,2p-1}^{(n)} \quad \forall \tilde{c}_m \in \tilde{\mathcal{C}}_p, \quad (32)$$

where

$$\theta_{k,2p}^{(i)} \triangleq \prod_{l=1}^{p-1} e^{(2\bar{b}_{2l}^{(i)} - 1) \frac{L_{2l}(k)}{2}} \quad \text{and} \quad \theta_{k,2p-1}^{(n)} \triangleq \prod_{l=1}^{p-1} e^{(2\bar{b}_{2l-1}^{(n)} - 1) \frac{L_{2l-1}(k)}{2}}.$$

Rewriting (26) as a double sum over  $i$  and  $n$  with  $\tilde{c}_m = [2i-1]d_p + j[2n-1]d_p$  and using the result of (32), it follows that:

$$\begin{aligned} \bar{\Omega}_k(\tau) &= 4\beta_k \times \\ &\sum_{i=1}^{2^{p-1}} \sum_{n=1}^{2^{p-1}} \left[ \theta_{k,2p}^{(i)} e^{-\rho [2i-1]^2 d_p^2} \cosh \left( \frac{\sqrt{E_s} [2i-1] d_p}{\sigma^2} u_k(\tau) + \frac{L_{2p}(k)}{2} \right) \right. \\ &\left. \times \theta_{k,2p-1}^{(n)} e^{-\rho [2n-1]^2 d_p^2} \cosh \left( \frac{\sqrt{E_s} [2n-1] d_p}{\sigma^2} v_k(\tau) + \frac{L_{2p-1}(k)}{2} \right) \right]. \end{aligned}$$

Then, after splitting the two sums, we obtain the following much useful factorization of  $\bar{\Omega}_k(\tau)$ :

$$\bar{\Omega}_k(\tau) = 4\beta_k F_{k,2p}(u_k(\tau)) F_{k,2p-1}(v_k(\tau)), \quad (33)$$

where

$$F_{k,q}(x) = \sum_{i=1}^{2^{p-1}} \theta_{k,q}^{(i)} e^{-\rho [2i-1]^2 d_p^2} \cosh \left( \frac{\sqrt{E_s} [2i-1] d_p}{\sigma^2} x + \frac{L_q(k)}{2} \right), \quad (34)$$

in which  $q$  is a generic counter that is used from now on to refer to  $2p$  or  $2p-1$  depending on the context. Finally, by using (33) back in (24) and dropping the constant terms, the LLF develops into<sup>4</sup>:

$$\mathcal{L}(\mathbf{y}; \tau) = \sum_{k=1}^K \ln \left( F_{k,2p}(u_k(\tau)) \right) + \sum_{k=1}^K \ln \left( F_{k,2p-1}(v_k(\tau)) \right). \quad (35)$$

<sup>4</sup>Note here that equivalent interesting decompositions for the LLF have been recently used to derive closed-form CA CRLBs for the code-aided estimation of other key channel parameters from coded square-QAM signals, such as the SNR [5, 6], the phase and frequency offsets [7, 8], and the DOA [9].

We succeeded here in decomposing the LLF into two analogous terms [the two sums in (35)] involving each either RVs  $u_k(\tau)$  or  $v_k(\tau)$  that will be shortly shown to have almost the same distributions. This is actually the cornerstone result upon which we will establish in the next section the analytical expressions for the CA TD CRLBs.

#### IV. DERIVATION OF THE CA CRLBS

As an overall benchmark, the CRLB lower bounds the variance of any unbiased estimator,  $\hat{\tau}$ , of the time delay parameter, i.e.,  $\mathbb{E}\{(\hat{\tau} - \tau)^2\} \geq \text{CRLB}(\tau)$ . It is explicitly given by [19]:

$$\text{CRLB}(\tau) = \frac{1}{I(\tau)}, \quad (36)$$

where  $I(\tau)$  is the Fisher information for the received data:

$$I(\tau) = -\mathbb{E} \left\{ \frac{\partial^2 \mathcal{L}(\mathbf{y}; \tau)}{\partial \tau^2} \right\}. \quad (37)$$

Using (35) in (37) and owing to the linearity of the partial derivative and expectation operators, it immediately follows that:

$$I(\tau) = \sum_{k=1}^K [\gamma_{k,2p}(\tau) + \gamma_{k,2p-1}(\tau)], \quad (38)$$

where

$$\gamma_{k,2p}(\tau) \triangleq -\mathbb{E} \left\{ \partial^2 \ln(F_{k,2p}(u_k(\tau))) / \partial \tau^2 \right\}, \quad (39)$$

$$\gamma_{k,2p-1}(\tau) \triangleq -\mathbb{E} \left\{ \partial^2 \ln(F_{k,2p-1}(v_k(\tau))) / \partial \tau^2 \right\}. \quad (40)$$

In addition, after some algebraic manipulations<sup>5</sup>, we show that:

**LEMMA 2:**  $u_k(\tau)$  and  $v_k(\tau)$  are two *independent* RVs whose distributions are given by:

$$p[u_k(\tau)] = \frac{2\beta_{k,2p}}{\sqrt{2\pi\sigma^2}} F_{k,2p}(u_k(\tau)) e^{-\frac{u_k(\tau)^2}{2\sigma^2}}, \quad (43)$$

$$p[v_k(\tau)] = \frac{2\beta_{k,2p-1}}{\sqrt{2\pi\sigma^2}} F_{k,2p-1}(v_k(\tau)) e^{-\frac{v_k(\tau)^2}{2\sigma^2}}. \quad (44)$$

As seen from (43) and (44), the two RVs  $u_k(\tau)$  and  $v_k(\tau)$  are *almost* identically distributed (i.e., their pdfs have the same structure, but they are parameterized differently). Therefore, when evaluating the required expectation with respect to either  $u_k(\tau)$  or  $v_k(\tau)$ , equivalent derivation steps can be followed to find either  $\gamma_{k,2p}(\tau)$  or  $\gamma_{k,2p-1}(\tau)$ . As such, we will only derive  $\gamma_{k,2p}(\tau)$  and later deduce the expression of  $\gamma_{k,2p-1}(\tau)$  by easy identification. To that end, we denote the first and second derivatives of  $F_{k,2p}(x)$  with respect to the working variable  $x$  in (34) by  $F'_{k,2p}(x)$  and  $F''_{k,2p}(x)$ , respectively. Using these notations, we show that:

$$\begin{aligned} & \frac{\partial^2}{\partial \tau^2} \ln(F_{k,2p}(u_k(\tau))) \\ &= \dot{u}_k^2(\tau) \left[ \frac{F''_{k,2p}(u_k(\tau))}{F_{k,2p}(u_k(\tau))} - \frac{F'_{k,2p}(u_k(\tau))^2}{F_{k,2p}(u_k(\tau))^2} \right] + \ddot{u}_k(\tau) \frac{F'_{k,2p}(u_k(\tau))}{F_{k,2p}(u_k(\tau))}, \end{aligned}$$

in which  $\dot{u}_k(\tau) \triangleq \partial u_k(\tau) / \partial \tau$  and  $\ddot{u}_k(\tau) \triangleq \partial^2 u_k(\tau) / \partial \tau^2$ . Moreover, using the same rationale as in Appendix C of [12], it can be easily shown that  $\dot{u}_k(\tau)$  and  $u_k(\tau)$  are two independent RVs. Thus, by applying the expectation operator to the previous equation, we obtain  $\gamma_{k,2p}(\tau)$  as follows:

$$\begin{aligned} \gamma_{k,2p}(\tau) &= \mathbb{E} \left\{ \dot{u}_k^2(\tau) \right\} \left[ \mathbb{E} \left\{ \frac{F''_{k,2p}(u_k(\tau))}{F_{k,2p}(u_k(\tau))} \right\} - \mathbb{E} \left\{ \frac{F'_{k,2p}(u_k(\tau))^2}{F_{k,2p}(u_k(\tau))^2} \right\} \right] \\ &\quad - \mathbb{E} \left\{ \ddot{u}_k(\tau) \frac{F'_{k,2p}(u_k(\tau))}{F_{k,2p}(u_k(\tau))} \right\}. \quad (45) \end{aligned}$$

<sup>5</sup>details were omitted due to space limitations

Since the pdfs of these RVs were already established in (43) and (44), the expectations involved above can be expressed in closed form by integration over  $p[u_k(\tau)]$  and  $p[v_k(\tau)]$ . For instance, we have:

$$\begin{aligned} \mathbb{E} \left\{ \left( \frac{F'_{k,2p}(u_k(\tau))}{F_{k,2p}(u_k(\tau))} \right)^2 \right\} &= \int_{\mathbb{R}} \frac{F_{k,2p}'^2(u_k(\tau))}{F_{k,2p}^2(u_k(\tau))} p[u_k(\tau)] du_k(\tau) \\ &= \frac{2\beta_{k,2p}}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} \frac{F_{k,2p}'^2(u_k(\tau))}{F_{k,2p}^2(u_k(\tau))} e^{-\frac{u_k(\tau)^2}{2\sigma^2}} du_k(\tau). \end{aligned}$$

After using the explicit expression of  $F'_{k,2p}(u_k(\tau))$ , the last equality is further simplified by using the variable substitution  $t = \sqrt{2}u_k(\tau)/\sigma$  to obtain:

$$\mathbb{E} \left\{ \left( \frac{F'_{k,2p}(u_k(\tau))}{F_{k,2p}(u_k(\tau))} \right)^2 \right\} = \frac{2\rho}{\sigma^2} \Psi_{k,2p}(\rho), \quad (46)$$

where  $\Psi_{k,2p}(\cdot)$  in the last equality is given by:

$$\Psi_{k,2p}(\rho) = \frac{\beta_{k,2p} d_p^2}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \frac{\lambda_{k,2p}^2(t, \rho)}{\delta_{k,2p}(t, \rho)} e^{-t^2/4} dt, \quad (47)$$

with

$$\begin{aligned} \lambda_{k,2p}(t, \rho) &= \sum_{i=1}^{2^{p-1}} (2i-1) \theta_{k,2p}^{(i)} e^{-[2i-1]^2 d_p^2 \rho} \\ &\quad \times \sinh \left( \sqrt{\rho} [2i-1] d_p t + \frac{L_{2p}(k)}{2} \right), \end{aligned}$$

$$\delta_{k,2p}(t, \rho) = \sum_{i=1}^{2^{p-1}} \theta_{k,2p}^{(i)} e^{-[2i-1]^2 d_p^2 \rho} \cosh \left( \sqrt{\rho} [2i-1] d_p t + \frac{L_{2p}(k)}{2} \right).$$

Using tedious algebraic manipulation, the remaining expectations in (45) are evaluated in closed form and the final results are giving by (for detailed derivations see [21]):

$$\begin{aligned} \mathbb{E} \left\{ \dot{u}_k(\tau)^2 \right\} &= E_s \sum_{l=1}^K \left( \omega_{l,2p} - \alpha_{l,2p}^2 \right) \dot{g}^2([l-k]T) \\ &\quad + E_s \left( \sum_{l=1}^K \alpha_{l,2p} \dot{g}([l-k]T) \right)^2 - \sigma^2 \dot{g}(0). \quad (48) \end{aligned}$$

$$\begin{aligned} \mathbb{E} \left\{ \ddot{u}_k(\tau) \frac{F'_{k,2p}(u_k(\tau))}{F_{k,2p}(u_k(\tau))} \right\} \\ = 2\rho \left[ \omega_{k,2p} \ddot{g}(0) + \alpha_{k,2p} \sum_{l \neq k} \alpha_{l,2p} \ddot{g}([l-k]T) \right], \quad (49) \end{aligned}$$

$$\mathbb{E} \left\{ \frac{F''_{k,2p}(u_k(\tau))}{F_{k,2p}(u_k(\tau))} \right\} = \frac{2\omega_{k,2p}}{\sigma^2} \rho. \quad (50)$$

with

$$\omega_{k,q} \triangleq 2\beta_{k,q} \cosh \left( \frac{L_q(k)}{2} \right) \sum_{i=1}^{2^{p-1}} \theta_{k,q}^{(i)} d_p^2 (2i-1)^2, \quad (51)$$

$$\alpha_{k,q} \triangleq 2\beta_{k,q} \sinh \left( \frac{L_q(k)}{2} \right) \sum_{i=1}^{2^{p-1}} \theta_{k,q}^{(i)} d_p (2i-1). \quad (52)$$

Finally, by injecting (48), (46), (50), and (49) back into (45), the analytical expression of  $\gamma_{k,2p}(\tau)$  is obtained in (41) given at the top of the current page. Due to the apparent symmetries between the distributions of the two RVs  $u_k(\tau)$  and  $v_k(\tau)$ , the analytical expression of  $\gamma_{k,2p-1}(\tau)$  can be directly deduced from the one of  $\gamma_{k,2p}(\tau)$  by easy identifications as given by (42) displayed on the top of the next page as well. The closed-form expression for

$$\begin{aligned} \gamma_{k,2p}(\tau) = 4\rho^2 \left[ \omega_{k,2p} - \Psi_{k,2p}(\rho) \right] & \left[ \sum_{l=1}^K \left( \omega_{l,2p} - \alpha_{l,2p}^2 \right) \dot{g}^2([l-k]T) + \left( \sum_{l=1}^K \alpha_{l,2p} \dot{g}([l-k]T) \right)^2 \right] \\ & - 2\rho \left[ \Psi_{k,2p}(\rho) \ddot{g}(0) - \alpha_{k,2p} \sum_{l \neq k} \alpha_{l,2p} \ddot{g}([l-k]T) \right]. \end{aligned} \quad (41)$$

$$\begin{aligned} \gamma_{k,2p-1}(\tau) = 4\rho^2 \left[ \omega_{k,2p-1} - \Psi_{k,2p-1}(\rho) \right] & \left[ \sum_{l=1}^K \left( \omega_{l,2p-1} - \alpha_{l,2p-1}^2 \right) \dot{g}^2([l-k]T) + \left( \sum_{l=1}^K \alpha_{l,2p-1} \dot{g}([l-k]T) \right)^2 \right] \\ & - 2\rho \left[ \Psi_{k,2p-1}(\rho) \ddot{g}(0) + \alpha_{k,2p-1} \sum_{l \neq k} \alpha_{l,2p-1} \ddot{g}([l-k]T) \right]. \end{aligned} \quad (42)$$

the TD CA CRLB is then obtained as the inverse of the Fisher information given by (38), i.e.:

$$\text{CRLB}(\tau) = \frac{1}{\sum_{k=1}^K \gamma_{k,2p}(\tau) + \gamma_{k,2p-1}(\tau)}. \quad (53)$$

#### A. Evaluation of the analytical CA CRLBs

In order to compute and plot the new CA CRLBs, one needs to evaluate the coefficients  $\omega_{k,q}$ ,  $\alpha_{k,q}$ , and  $\Psi_{k,q}(\rho)$  that are involved in the quantities  $\gamma_{k,q}(\tau)$  for  $q = 2p$  and  $q = 2p - 1$ . These coefficients are, however, functions of the *a priori* LLRs,  $L_l(k)$ , as seen from (47), (51) and (52), respectively. In the sequel, we briefly explain how these LLRs can be obtained from the output of the SISO decoders at the convergence of the BCJR algorithm. First, the matched filter returns a sequence of  $K$  symbol-rate samples:

$$\mathbf{y}(\tau) = [y_1(\tau), y_2(\tau), \dots, y_K(\tau)]^T, \quad (54)$$

where

$$y_k(\tau) = \int_{-\infty}^{+\infty} y(t)h(t - kT - \tau)dt = \sqrt{E_s} a(k) + w_k(\tau). \quad (55)$$

Then, the soft demapper extracts the so-called *bit likelihoods* [20]:

$$\Lambda_l(k) \triangleq \ln \left( \frac{p[\mathbf{y}(\tau)|b_l^k = 1]}{p[\mathbf{y}(\tau)|b_l^k = 0]} \right), \quad (56)$$

for all the code bits and feed them as inputs to the turbo decoder. By exchanging the so-called *extrinsic* information between the two SISO decoders, the *a posteriori* LLRs of the code bits:

$$\Upsilon_l(k) = \ln \left( \frac{P[b_l^k = 1|\mathbf{y}(\tau)]}{P[b_l^k = 0|\mathbf{y}(\tau)]} \right). \quad (57)$$

are updated iteratively according to the turbo principle. We denote their values at the  $r^{\text{th}}$  turbo iteration as  $\Upsilon_l^{(r)}(k)$ . After say  $R$  turbo iterations, a steady state is achieved wherein  $\Upsilon_l^{(R)}(k) \approx \Upsilon_l(k)$ , for every  $l$  and  $k$ , and their signs are used to detect the bits. Yet, owing to the well-known Bayes' formula, we have:

$$P[b_l^k = 1|\mathbf{y}(\tau)] = \frac{p[\mathbf{y}(\tau)|b_l^k = 1]P[b_l^k = 1]}{p[\mathbf{y}(\tau)]}, \quad (58)$$

and

$$P[b_l^k = 0|\mathbf{y}(\tau)] = \frac{p[\mathbf{y}(\tau)|b_l^k = 0]P[b_l^k = 0]}{p[\mathbf{y}(\tau)]}. \quad (59)$$

Therefore, by taking the ratio of (58) and (59) and applying the natural logarithm, it immediately follows that:

$$L_l(k) = \Upsilon_l(k) - \Lambda_l(k) \approx \Upsilon_l^{(R)}(k) - \Lambda_l(k), \quad (60)$$

meaning that the required *a priori* LLRs of the code bits can be easily obtained from their steady-state *a posteriori* LLRs and  $\Lambda_l(k)$  already computed by the *soft* demapper prior to data decoding.

## V. SIMULATION RESULTS

In this section, we provide some graphical representations for the new TD CA CRLBs for different modulation orders and different coding rates. The encoder is composed of two identical RSCs concatenated in parallel, having generator polynomials (1,0,1,1) and (1,1,0,1), and a systematic rate  $R_0 = \frac{1}{2}$  each. The output of the turbo encoder is punctured in order to achieve the desired code rate  $R$ . For the tailing bits, the size of the RSC encoders memory is fixed to 4. We consider a root-raised-cosine (RRC) signal with rolloff factor  $\alpha = 0.2$ . We also consider QPSK and 16-QAM, as two representative examples of square-QAM constellations, and two different coding rates, namely  $R = \frac{1}{2}$  and  $R = \frac{1}{3}$ . We begin by verifying in Figs. 2 and 3 that the new analytical CA CRLBs coincide with their *empirical* counterparts obtained previously in [10] from exhaustive Monte-Carlo simulations<sup>6</sup>. In fact, unlike our closed-form solution, an extremely large number of noisy observations was generated in [10] in order to find an empirical value for the expectation involved in the Fisher information (37). Hence, our new analytical expression corroborates these previous attempts to evaluate the underlying TD CA CRLBs *empirically* and allow their immediate evaluation for any square-QAM turbo-coded signal<sup>7</sup>.

As expected, we also see from both figures that the CA CRLBs are smaller than their NDA counterparts. This highlights the performance improvements that can be achieved by a coded system (over an uncoded one) by exploiting the information about the transmitted bits that is obtained from the SISO decoders. Additionally and most prominently, the CA CRLBs decrease rapidly and reach the DA CRLBs which are the best bounds ever one would be able to achieve if all the transmitted symbols were perfectly known to the receiver, hypothetically.

<sup>6</sup>For cross validation purposes, note that during our simulations the steady-state *a posteriori* probabilities of the symbols required by the empirical approach of [10] were computed from their steady-state *a priori* probabilities using Bayes' formula. The latter are obtained as the product of the steady-state *a priori* probabilities of the corresponding conveyed bits which are themselves obtained from the steady-state *a priori* LLRs in (60) using (9).

<sup>7</sup>Note also that our analytical CA CRLBs are averaged over a very small number of realizations (20 in our simulations) in order to smoothen the corresponding curves.

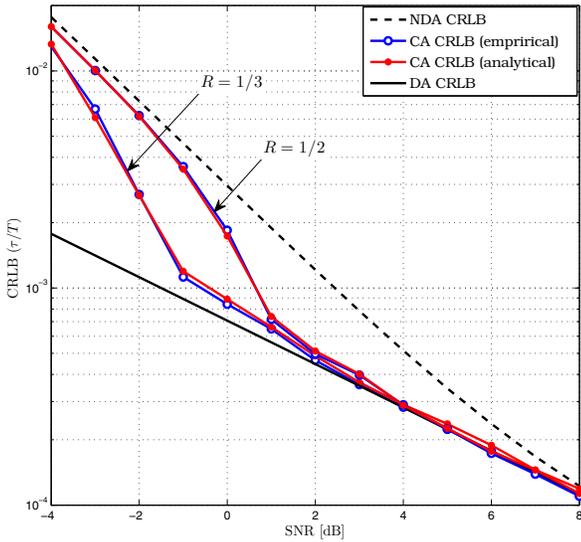


Fig. 2. Comparison between the empirical and analytical CA CRLBs for different code rates as function of the SNR: QPSK, rolloff = 0.2.

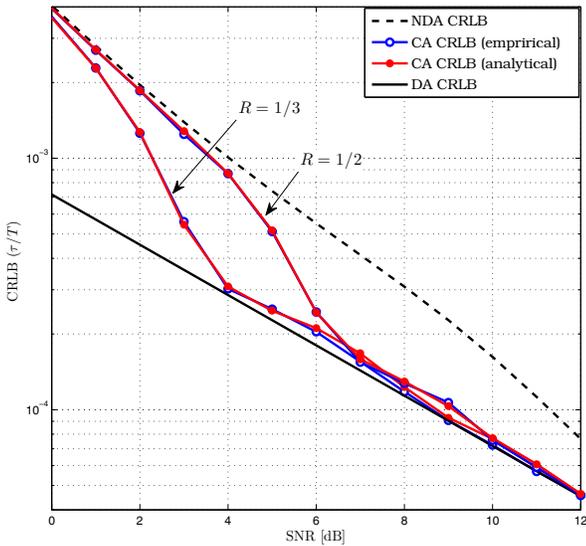


Fig. 3. Comparison between the empirical and analytical CA CRLBs for different code rates as function of the SNR: 16-QAM, rolloff = 0.2.

## VI. CONCLUSION

In this paper, we derived for the first time the closed-form expressions of the Cramér-Rao lower bounds for code-aided symbol timing estimation from turbo-coded square-QAM transmissions. The new CA CRLBs revealed the huge performance improvements in terms of timing recovery that can be achieved by exploiting the soft information delivered by SISO decoders at each turbo iteration. The new analytical CRLBs coincide exactly with their empirical counterparts established in previous pioneering works on the subject exhaustive Monte-Carlo simulations.

## REFERENCES

- [1] 3GPP TS 36.211: 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation.
- [2] 3GPP TS 36.212 v8.0.0 (2007-09): "Multiplexing and channel coding (FDD) (Release 8)". Available online at: <http://www.3gpp.org/Highlights/LTE/LTE.htm>

- [3] U. Mengali and A. N. D'andrea, *Synchronization Techniques for Digital Receivers*, New York: Plenum, 1997.
- [4] C. Berrou, "The ten-year old turbo codes are entering into service", *IEEE J. Commun. Mag.*, vol. 41, no. 8, pp. 110-116, Aug. 2003.
- [5] F. Bellili, A. Methenni, and S. Affes, "Closed-form CRLBs for SNR estimation from turbo-coded square-QAM-modulated signals," in *Proc. of IEEE Globecom*, Austin, TX, USA, Dec. 8-12, 2014, pp. 1765-1771.
- [6] F. Bellili, A. Methenni, and S. Affes, "Closed-form CRLBs for SNR estimation from turbo-coded BPSK-, MSK-, and square-QAM-modulated signals," *IEEE Trans. Signal Process.*, vol. 62, no. 15, pp. 4018-4033, Aug. 2014.
- [7] F. Bellili, A. Methenni, and S. Affes, "Closed-form Cramér-Rao lower bounds for CFO and phase estimation from turbo-coded square-QAM-modulated signals," in *Proc. of IEEE Globecom*, Austin, TX, USA, Dec. 8-12, 2014, pp. 4083-4089.
- [8] F. Bellili, A. Methenni, and S. Affes, "Closed-form CRLBs for CFO and phase estimation from turbo-coded square-QAM-modulated transmissions," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2513-2531, May 2015.
- [9] F. Bellili, C. Elguet, S. Ben Amor, S. Affes, and A. Stéphane, "Closed-form Cramer-Rao lower bounds for DOA estimation from turbo-coded square-QAM-modulated transmissions," in *Proc. of IEEE ICASSP*, South Brisbane, QLD, Australia, Apr. 19-24, 2015, pp. 3492-3496.
- [10] N. Noels, H. Wymeersch, H. Steendam, and M. Moeneclaey, "Carrier and clock recovery in (turbo) coded systems: Cramer-Rao bound and synchronizer performance," *EURASIP J. Appl. Signal Process.*, vol. 2005, no. 6, pp. 972-980, May 2005.
- [11] N. Noels, H. Wymeersch, H. Steendam, and M. Moeneclaey, "True Cramér-Rao bound for timing recovery from a bandlimited linearly modulated waveform with unknown carrier phase and frequency," *IEEE Trans. Comm.*, vol. 52, no. 3, pp. 473-383, Mar. 2004.
- [12] A. Masmoudi, F. Bellili, S. Affes, and A. Stéphane, "Closed-form expressions for the exact Cramér-Rao bounds of timing recovery estimators from BPSK, MSK and square-QAM transmissions," *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2474-2484, June 2011.
- [13] J. R. Barry, A. Kavcic, S. W. McLaughlin, A. Nayak, and W. Zeng, "Iterative timing recovery," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 89-102, Jan. 2004.
- [14] C. Herzet, V. Ramon, and L. Vandendorpe, "Turbo synchronization: a combined sum-product and expectation-maximization algorithm approach," in *Proc. IEEE SPAWC'05*, New-York, USA, June 5-8, 2005, pp. 191-195.
- [15] C. Herzet, H. Wymeersch, M. Moeneclaey, and L. Vandendorpe, "On maximum-likelihood timing synchronization," *IEEE Trans. Commun.*, vol. 55, no. 6, pp. 1116-1119, June 2007.
- [16] N. Wu, H. Wang, J. M. Kuang, and C. X. Yan, "Performance analysis of code-aided symbol timing recovery on AWGN channels," *IEEE Trans. Commun.*, vol. 59, no. 7, pp. 1975-1984, Jul. 2011.
- [17] C. Herzet, V. Ramon, and L. Vandendorpe, "A theoretical framework for iterative synchronization based on the sum-product and the expectation-maximization algorithms," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1644-1658, May 2007.
- [18] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate", *IEEE Trans. Inf. Theory*, vol. IT-20(2), no. 2, pp. 284-287, Mar. 1974.
- [19] S.M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Upper Saddle River, NJ: Prentice Hall, 1993.
- [20] S. L. Goff, A. Glavieux, and C. Berrou, "Turbo-codes and high spectral efficiency modulation," in *Proc. of IEEE ICC'94*, New Orleans, LA, USA, May 1994, pp. 645-649.
- [21] F. Bellili, A. Methenni, S. Ben Amor, S. Affes, and A. Stéphane, "Time synchronization of turbo-coded square-QAM-modulated transmissions: Code-aided ML estimator and closed-form Cramér-Rao lower bounds," available online at <http://arxiv.org/abs/1509.03810>.