Code-Aided DOA Estimation From Turbo-Coded QAM Transmissions: Analytical CRLBs and Maximum Likelihood Estimator

Faouzi Bellili, Chaima Elguet, Souheib Ben Amor, Sofiène Affes, *Senior Member, IEEE*, and Alex Stéphenne, *Senior Member, IEEE*

***Abstract*— In this paper, we address the problem of direction of arrival (DOA) estimation from turbo-coded square-QAM-modulated transmissions. We devise a new code-aware direction finding concept, derived from maximum likelihood (ML) theory, wherein the soft information provided by the soft-input soft-output decoder, in the form of log-likelihood ratios, is exploited to assist the estimation process. At each turbo iteration, the decoder output is used to refine the ML DOA estimate. The latter is in turn used to perform a more focused receiving beamforming thereby providing more reliable information-bearing sequences for the next turbo iteration. In order to benchmark the new estimator, we also derive the analytical expressions for the exact Cramer-Rao lower bounds (CRLBs) of code-aided (CA) DOA estimates. Simulation results will show that the new CA direction finding scheme lies between the two traditional schemes of completely non-data-aided and data-aided (DA) estimations. Huge performance improvements are achieved by embedding the direction finding and receive beamforming tasks into the turbo iteration loop. Moreover, the new CA DOA estimator reaches the new CA CRLBs over a wide range of practical SNRs thereby confirming its statistical efficiency. As expected intuitively, its performance further improves at higher coding rates and/or lower modulation orders.**

***Index Terms*— Turbo-codes, direction of arrival estimation, beamforming, log-likelihood ratio, maximum likelihood, Cramér-Rao lower bound, QAM signals, non-data-aided, data-aided and code-aided estimation.**

I. Introduction

THE rapid development of radio communications has resulted in a growing demand for location-aware services fueled in part by the soaring need to increasingly accommodate users' mobility. Clearly, highly accurate DOA estimates are crucial for enhanced overall system performance. However, grasping this high level of estimation accuracy requires a sufficiently broad theoretical foundation for the underlying direction finding technique and more than ever the ability to be grounded in practical situations. Actually, The problem of DOA estimation has been a hot array signal processing research topic over the last few decades and a suitable DOA estimator can be selected from a plethora of state-of-the-art techniques (see [2], [3] and references therein). Recently, there has been a resurgence of interest in developing advanced DOA estimators that are specifically tailored to the emergent massive MIMO systems. Interested readers are referred to the very recent works in [4], [5] which introduce novel 2-D DOA estimation techniques that are geared toward massive MIMO systems in presence of multiple incoherently distributed sources. Depending on how the recorded data are processed to output the required DOA estimate, DOA estimation techniques can be broadly categorized into two major categories: i) subspace (SS)-based or ii) ML-based methods. To their credentials, SS-based approaches are known to be computationally less expensive. However, as they extract the DOA information from the covariance matrix of the received data instead of the data themselves, they are usually suboptimal [6]. They hence suffer from severe performance degradation at low SNR levels and/or small numbers of snapshots. ML approaches, however, apply the estimation process directly on the received samples and always enjoy higher accuracy and enhanced resolution capabilities, however, very often at the cost of substantial increase in complexity [7].

Depending on the degree of *a priori* knowledge about the transmitted symbols, DOA estimators can be alternatively categorized into two other broad subcategories: NDA and DA approaches. NDA methods are *completely blind* and, as such, do also suffer from severe performance degradation in harsh SNR conditions. DA methods, however, are known to be highly accurate, but they impinge on the whole throughput of the system as they require the frequent insertion and transmission of perfectly known (i.e., pilot) sequences. It sounds reasonable then to conceive a third middle-ground alternative between these two extreme cases. In fact, rather than relying on perfectly known or completely unknown symbols, CA estimation takes advantage of the *soft information* delivered by the decoder at each turbo iteration. In plain English, the decoder assistance is called upon in an attempt to enhance the estimation performance yet with no impact on the spectral efficiency of the system. Spectral efficiency is, indeed, of paramount importance to current- and next-generation wireless communication systems which are compelled to provide high quality of service, while satisfying at the same time the ever-increasing demand for high data rates and capacity. In order to meet these requirements, it is also advocated to use powerful error-correcting codes in conjunction with high-spectral-efficiency

Manuscript received August 13, 2015; revised April 10, 2016, August 16, 2016, and January 3, 2017; accepted January 20, 2017. Date of publication March 17, 2017; date of current version May 8, 2017. This work was supported by a Canada Research Chair in Wireless Communications and a Discovery Accelerator Supplement from NSERC. This paper was presented at IEEE ICASSP, South Brisbane, QLD, Australia, April 2015 [1]. The associate editor coordinating the review of this paper and approving it for publication was L. Liu.

The authors are with the INRS-EMT, Montréal, QC H5A, Canada (e-mail: bellili@emt.inrs.ca; chaima.elguet@emt.inrs.ca; souheib.ben.amor@emt.inrs.ca; affes@emt.inrs.ca; stephenne@ieee.org).

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

Digital Object Identifier 10.1109/TWC.2017.2669026

modulations such as high-order QAMs. In fact, turbo codes along with higher-order QAMs are currently being adopted in Mobile WIMAX [8], long-term-evolution (LTE) and LTE advanced (LTE-A) or beyond (LTE-B) systems [9].

Turbo codes, in particular, surged as a promising choice for reliable high-data-rate communications when Berrou and Glavieux showed in [10] the possibility of transmitting data with code rates above the channel cut-off rate. Yet, beyond the simple introduction of a new error-correcting code, the turbo principle itself has since then sparked a new era in the theory and practices of communication engineering. It opened up a whole new way of thinking regarding the design of communication and signal processing algorithms. In this context, the idea of relying on the decoder assistance has been heavily investigated in several research works and lead to many advanced concepts including the so-called turbo equalization, turbo synchronization, turbo channel estimation and detection, etc. Yet, a common principle to all the aforementioned concepts is to exploit the decoder's output via iterative feedback so as to enhance the estimates of the various involved channel parameters (i.e., CA estimation). The so-called decoder's assistance has indeed resulted in dramatic performance enhancements within the context of turbo equalization of frequency-selective channels (see [11] and references therein). A plethora of research works dealing with the CA estimation of phase, frequency, and time offsets within the framework of turbo synchronization have also been reported in the open literature (see [12]–[26] and references therein). The feedback of the decoder was also leveraged in order to perform turbo SNR estimation and decoding [27], [28], turbo channel estimation and decoding [29]–[31] as well as turbo detection and decoding [32].

To the best of the authors' knowledge, however, no contribution has so far dealt with the problem of CA DOA estimation. Actually, an exhaustive bibliographical review revealed that very few works (cf. [33]–[40]) have so far tackled the problem of DOA estimation in presence of *modulated* sources. Indeed, most of the existing techniques assume the unknown signal to be either deterministic or random but Gaussian distributed leading, respectively, to the *conditional* [41]–[45] or *unconditional* DOA ML estimators [46]–[50]. Yet, it has been recently shown in [51] and [52] that the Gaussian assumption leads to the largest CRLBs, i.e., the associated bounds underestimate the actual achievable performance. In fact, the Gaussian assumption is a natural choice when nothing is known beforehand about the exact data distribution [51]. Yet, with the Gaussian distribution, the transmitted symbols are assumed to possibly take any complex value while they are actually drawn in digital transmissions from a finite-size set (namely, the constellation alphabet). Naturally, ignoring this fact — by simply resorting to the Gaussian assumption — introduces more ambiguity about the unknown symbols thereby leading to suboptimal estimation performances. In our work, however, the actual *a priori* probabilities (APPs) of the transmitted symbols (i.e., their actual distribution) are made available from the output of the SISO decoders and exploited during the direction finding process in order to enhance the estimation performance.

All these facts motivated us to tackle in this paper the problem of CA DOA estimation from higher-order square-QAM turbo-coded signals. We propose, indeed, a low-cost CA DOA estimator that is derived from ML theory wherein the turbo decoding and DOA estimation tasks operate side by side. In fact, at each turbo iteration, the proposed algorithm leverages the *soft information* provided by the SISO decoders in order to refine the DOA estimate. The latter is in turn used to enhance the decoding performance — during the next turbo iteration — by use of a more accurately steered receive beamformer. In order to appropriately benchmark the proposed estimator, we also derive the analytical expressions for the CRLBs of the underlying CA estimation problem. As a well-known fundamental lower bound, the CRLB quantifies the best achievable error variance in practice [53]. Its derivation is often considered as extremely challenging for higher-order modulations even under *uncoded* transmissions. Indeed, the very first DOA CRLBs for linearly-modulated sources appeared in 2006 [33] for *uncoded* BPSK and QPSK signals. It was only recently, though, that their analytical expressions have been extended to higher-order QAMs under *uncoded* transmissions as well [34]. In order to validate our new analytical procedure of evaluating the CA DOA CRLBs, we also evaluate the latter bounds using another approach that computes the underlying bounds empirically, i.e., from exhaustive Monte-Carlo simulations. The empirical approach was developed within the context of turbo synchronization and is adapted here to the direction finding problem. From the algorithmic point of view, we also verify that the new CA ML estimator achieves the derived CA CRLBs over a wide range of practical SNRs thereby confirming its statistical efficiency.

Note here that the existing CA estimation approaches can also be applied with appropriate modifications to the direction finding problem at hand. In particular, under the ML estimation umbrella, the combination of the sum-product (SP) and expectation-maximization (EM) concepts pioneered in [19] can be leveraged in order to find the CA ML estimates (MLEs) of the DOA parameter. Yet, as will be seen in Section VI, the complexity of the resulting SP-EM algorithm is very high compared to the proposed ML estimator. This is because, in the SP-EM-based ML approach, an EM iteration loop is required in each turbo iteration wherein the algorithm switches between the so-called expectation step (E-STEP) and maximization step (M-STEP). Indeed, in each turbo iteration, SP-EM performs the following main four steps to complete a single EM iteration:

- Obtain new beamforming outputs using the DOA estimate of the previous EM iteration;
- Update the symbols' *a posteriori* probabilities (APoPs) using the new beamforming output;
- Marginalize *empirically* the conditional (on the transmitted symbols) likelihood function (LF) with respect to those APoPs (E-STEP);
- Maximize the marginalized LF with respect to the working DOA variable (M-STEP).

At the convergence of the EM algorithm, the obtained DOA estimate is used to acquire more refined beamforming outputs which will serve as input for the next turbo

iteration where all the aforementioned EM-related steps are repeated.

In this paper, however, we exploit the full symmetry of square-QAM constellations in order to derive a an analytical expression for the log-likelihood function (LLF). In particular, marginalization of the conditional LF with respect to the transmitted symbols is carried out analytically and the *a priori* LLRs of the elementary code bits are explicitly incorporated in the LLF expression. We propose thereof a more systematic framework to their direct integration into the CA estimation process, thereby eliminating completely the need for an EM iteration loop under each turbo iteration. In other words, contrarily to SP-EM, the obtained LLF needs to be maximized only once per turbo iteration after being updated by the associated *a priori* LLRs. Consequently, the proposed CA joint DOA estimation and beamforming scheme offers significant improvements in computational complexity as compared to the existing SP-EM approach as will be seen later in Section VI. It also enjoys a remarkable performance advantage at low SNR thresholds and/or higher-order modulations.

The rest of this paper is organized as follows. In Sections II and III, respectively, we introduce the system model and derive the corresponding LLF. In Section IV, we derive the new analytical CA CRLBs and introduce the empirical approach developed for their validation. In Section V, we introduce the new CA ML DOA estimator. In section VI, we assess its performance and benchmark it against the new CA CRLBs as well as several state-of-the-art techniques. Finally, we draw out some concluding remarks in Section VII.

We mention beforehand that some of the common notations will be used throughout this paper. Vectors and matrices are represented by lower- and upper-case bold fonts, respectively. Moreover, $\{.\}^H$ denotes the Hermitian (i.e., transpose conjugate) of any vector or matrix. The operators $\Re\{.\}$ and $\Im\{.\}$ return the real and imaginary parts of any complex number, respectively. Moreover, $\{.\}^*$ and $|.|$ return its conjugate and its amplitude, respectively, and j is the pure complex number that verifies $j^2 = -1$. We will also denote the probability mass function (PMF) for discrete random variables (RVs) by $Pr[.]$ and the probability density function (pdf) for continuous RVs by $p[.]$. The statistical expectation is denoted as $\mathbb{E}\{.\}$ and the notation $\triangleq$ is used for definitions.

II. System Model

Consider a turbo-coded system consisting of two recursive systematic convolutional codes (RSCs) which are concatenated in parallel via an interleaver Π_1. The desired coding rate R is achieved by properly puncturing the parity bits out of each RSC encoder. The entire coded bit sequence is then scrambled with an outer interleaver, Π_2, and divided into K subgroups of $2q$ bits for some integer $q \geq 1$. The k^{th} subgroup of coded bits, $b_1^k b_2^k \cdots b_l^k \cdots b_{2q}^k$, is conveyed by a symbol $x(k)$ that is selected from a fixed alphabet, $\mathcal{C} = \{c_0, c_1, \cdots, c_{M-1}\}$, of a $M-$ary (with $M = 2^{2q}$) QAM constellation (i.e., square-QAM). In fact, each point, $c_m \in \mathcal{C}$, is mapped onto a unique sequence of $\log_2(M) = 2q$ bits denoted here as $\bar{b}_1^m \bar{b}_2^m \cdots \bar{b}_l^m \cdots \bar{b}_{2q}^m$, according to the Gray coding mechanism, and c_m is selected to convey the k^{th} subgroup of coded bits [i.e., $x(k) = c_m$] if and only if $b_l^k = \bar{b}_l^m$ for $l = 1, 2, \cdots, 2q$. We also define the so-called *a priori* LLR of the l^{th} coded bit, b_l^k, conveyed by the transmission of $x(k)$, as follows:

$$L_l(k) \triangleq \ln\left(\frac{Pr\left[b_l^k = 1\right]}{Pr\left[b_l^k = 0\right]}\right), \tag{1}$$

where $Pr\left[b_l^k = 1\right]$ and $Pr\left[b_l^k = 0\right]$ are the *a priori* probabilities (APPs) of the coded bit b_l^k. Exploiting (1) and the fact that $Pr\left[b_l^k = 0\right] + Pr\left[b_l^k = 1\right] = 1$, it immediately follows that the APPs of each conveyed bit are expressed in terms of its *a priori* LLR as follows:

$$Pr\left[b_l^k = 1\right] = \frac{e^{L_l(k)}}{1 + e^{L_l(k)}} \text{ and } Pr\left[b_l^k = 0\right] = \frac{1}{1 + e^{L_l(k)}}. \tag{2}$$

For mathematical convenience, the two identities in (2) will be merged together in the following single generic expression:

$$Pr[b_l^k = \bar{b}_l^m] = \frac{1}{2\cosh\left(L_l(k)/2\right)} e^{(\bar{b}_l^m - 1)\frac{L_l(k)}{2}}, \tag{3}$$

in which $\bar{b}_l^m$ can be either 0 or 1 depending on: i) the underlying Gray mapping rule, and ii) which of the symbols, $\{c_m\}_{m=1}^M$, is transmitted at time instant k.

The receiver is also equipped with N_a antenna branches and the received signal is corrupted by a complex circularly symmetric additive white Gaussian noise (AWGN) which is assumed to be temporally and spatially white. At each k^{th} time index, the noisy samples received over all the antenna elements are stacked into a single vector, $\mathbf{y}(k) = [y_1(k), y_2(k), \cdots, y_{N_a}(k)]^T$. In array signal processing terminology, $\mathbf{y}(k)$ is known as *snapshot* and it is modeled as follows [3]:

$$\mathbf{y}(k) = S\mathbf{a}(\theta)x(k) + \mathbf{w}(k), \text{ for } k = 0, 1, \ldots, K-1, \tag{4}$$

where S is the channel coefficient and $x(k)$ is the k^{th} unknown transmitted symbol. Moreover, $\mathbf{a}(\theta)$ is the steering vector which is a function of the unknown DOA, θ, to be estimated:

$$\mathbf{a}(\theta) = \left[e^{j\pi\varphi_1(\theta)}, e^{j\pi\varphi_2(\theta)}, \ldots, e^{j\pi\varphi_{N_a}(\theta)}\right]^T. \tag{5}$$

In (5), $\{\varphi_i(\theta)\}_{i=1}^{N_a}$ are some *real-valued* transformations (of the scalar DOA parameter) that depend on the geometry of the array configuration.[1] The random noise components, $\mathbf{w}(k)$, are modeled as complex circularly symmetric Gaussian vectors with independent real and imaginary parts and covariance matrix $\mathbb{E}\{\mathbf{w}(k)\mathbf{w}^H(k)\} = \sigma^2\mathbf{I}$. We also define the per-symbol SNR as follows:

$$\rho = \frac{S^2}{\sigma^2}. \tag{6}$$

Note here that the narrowband model in (4) is well justified in practice by its wide adoption in current and next-generation multicarrier communication systems, such as LTE and LTE-A or LTE-B. In fact, it is well known that OFDM systems

[1]For the explicit expressions of, $\{\varphi_i(\theta)\}_{i=1}^{N_a}$ under the widely studied uniform linear arrays (ULAs) and uniform circular arrays (UCAs), please refer to (51) in Section IV.

transform a multipath frequency-selective channel in the time domain into a frequency-flat (i.e., narrowband) channel over each subcarrier as modeled by (4) [54], [55]. Yet, even over traditional single-carrier systems, the narrowband model in (4) could still be valid in practice when the symbol duration is larger than the delay spread of the channel.

Moreover, as evidenced by (4), we tackle here the problem of DOA estimation of a single source. In practice, the single-source case is encountered, for instance, when dealing with CDMA signals after despreading. Indeed, one single source corresponding to the desired signal will be preserved through constructive correlation and any other source will be dramatically reduced by destructive correlation and incorporated in the noise component. The single-source model can also be obtained in the case of multisource transmissions if we apply an algorithm of blind source separation (BSS) as a post treatment. Hence, the observation obtained for each source will follow a single-source model over which our newly derived method can be adequately applied to estimate the DOA of the corresponding source.

III. Derivation of the LLF

The *global* LLF of the underlying estimation problem is defined as follows:

$$\mathcal{L}(\mathbf{Y};\theta) \triangleq \ln\big(p[\mathbf{Y};\theta]\big), \tag{7}$$

where $\mathbf{Y} = [\mathbf{y}(0)\ \mathbf{y}(1)\ \cdots\ \mathbf{y}(K-1)]$ is the space-time matrix that gathers all the recorded snapshots given in (4) and $p[\mathbf{Y};\theta]$ is the pdf of $\mathbf{Y}$ parameterized by θ. For estimation purposes only, a simplifying assumption is usually used under coded digital transmissions. This assumption postulates that the transmitted symbols are independent in spite of the statistical dependence introduced by channel coding. We emphasize, however, the fact that this simplifying assumption does not deny one to exploit the dependence of the coded bits during the decoding process itself. Indeed, the data dependence is exploited by the SISO decoders in order to output the estimates for the coded bits' *a posteriori* LLRs. The latter are then used to decode the bits and also to compute their *a priori* LLRs (as explained later in Section IV-B) which are in turn used to evaluate the CA CRLBs and to find the CA DOA ML estimate. Yet, even by assuming independent symbols, it turns out that no much information is lost from the estimation point of view. In fact, the performance of the resulting CA estimation scheme is equivalent to that of the ideal DA scheme (i.e., perfectly known symbols) over a wide range of practical SNRs. Using the assumption of independent symbols and recalling that the AWGN components are also assumed to be temporally white, the pdf $p[\mathbf{Y};\theta]$ in (7) can be factorized as:

$$p[\mathbf{Y};\theta] = \prod_{k=0}^{K-1} p[\mathbf{y}(k);\theta], \tag{8}$$

where the pdf of each snapshot, $p\big[\mathbf{y}(k);\theta\big]$, is given by:

$$p\big[\mathbf{y}(k);\theta\big] = \sum_{c_m \in C} Pr\big[x(k)=c_m\big]\ p\big[\mathbf{y}(k);\theta | x(k)=c_m\big]. \tag{9}$$

Here, $\big\{Pr[x(k)=c_m]\big\}_{c_m\in C}$ are the APPs of the transmitted symbols. Owing to the noise Gaussianity and using $\|\mathbf{a}(\theta)\|^2 = N_a$, it can be shown that:

$$p[\mathbf{y}(k);\theta] = \frac{1}{\pi^{N_a}\sigma^{2N_a}} \exp\left(\frac{1}{\sigma^2}||\mathbf{y}(k)||^2\right) D_k(\theta), \tag{10}$$

where $D_k(\theta)$ is explicitly given by:

$$D_k(\theta) = \sum_{c_m\in C} Pr[x(k)=c_m] \exp\left(-\frac{N_a S^2 |c_m|^2}{\sigma^2}\right) \times \exp\left(\frac{2}{\sigma^2}\Re\Big\{S\, c_m^* \mathbf{a}^H(\theta)\mathbf{y}(k)\Big\}\right). \tag{11}$$

By plugging (8)-(10) back into (7), the LLF of the underlying estimation problem develops into:

$$\mathcal{L}(\mathbf{Y};\theta) = -KN_a \ln\left(\pi\sigma^2\right) + \frac{1}{\sigma^2}\sum\nolimits_{k=0}^{K-1} ||\mathbf{y}(k)||^2 + \sum\nolimits_{k=0}^{K-1} \ln\big(D_k(\theta)\big). \tag{12}$$

Moreover, starting from (11) and resorting to tedious algebraic manipulations, equivalent to those recently used in [28], it can be shown that the term $D_k(\theta)$ can be factorized as follows:

$$D_k(\theta) = 4\beta_k\, H_{k,2q}\big(u_k(\theta)\big) H_{k,2q-1}\big(v_k(\theta)\big), \tag{13}$$

where the two functions $H_{k,2q}(.)$ and $H_{k,2q-1}(.)$ are commonly defined as follows:

$$H_{k,s}(x) = \sum_{i=1}^{2^{q-1}} \eta_{k,s}(i) e^{-\rho N_a [2i-1]^2 d_q^2} \times \cosh\left(\frac{2S[2i-1]\sqrt{N_a}d_q}{\sigma^2}x + \frac{L_s(k)}{2}\right),$$

with $2d_q$ being the inter-symbol distance, s a generic counter used from now on to refer either to $2q$ or $2q-1$ (depending on the context). Moreover, the terms $\eta_{k,2q}(i)$, $\eta_{k,2q-1}(n)$, and β_k are expressed as follows:

$$\eta_{k,2p}(i) = \prod_{l=1}^{q-1} e^{(2\bar{b}_{2l}^{(i)}-1)\frac{L_{2l}(k)}{2}}, \tag{14}$$

$$\eta_{k,2p-1}(n) = \prod_{l=1}^{q-1} e^{(2\bar{b}_{2l-1}^{(n)}-1)\frac{L_{2l-1}(k)}{2}}. \tag{15}$$

$$\beta_k = \prod_{l=1}^{2q} \frac{1}{2\cosh\big(L_l(k)/2\big)}. \tag{16}$$

Besides, $u_k(\theta)$ and $v_k(\theta)$ involved in (13) are, respectively, the real and imaginary parts of a receive-beamformer output steered toward the direction of the unknown DOA to be estimated:

$$u_k(\theta) = \frac{1}{\sqrt{N_a}}\Re\Big\{\mathbf{a}^H(\theta)\mathbf{y}(k)\Big\}$$
$$v_k(\theta) = \frac{1}{\sqrt{N_a}}\Im\Big\{\mathbf{a}^H(\theta)\mathbf{y}(k)\Big\}. \tag{17}$$

Using (13) in (12) and dropping the terms that do not depend on θ, it follows that the useful LLF:

$$\mathcal{L}(\mathbf{Y};\theta) = \sum_{k=0}^{K-1} \ln\Big(H_{k,2q}\big(u_k(\theta)\big)\Big) + \sum_{k=0}^{K-1} \ln\Big(H_{k,2q-1}\big(v_k(\theta)\big)\Big), \tag{18}$$

involves the sum of two analogous terms. In the next section, we will build upon this observation to derive for the first time new analytical expressions of the DOA CA CRLBs.

IV. New Analytical and Empirical CRLBs for CA DOA Estimation

A. Derivation of the Analytical CA CRLBs

The CRLB lower bounds the variance of any unbiased estimator $\widehat{\theta}$ of θ, i.e., $\mathbb{E}\{(\widehat{\theta}-\theta)^2\} \geq \text{CRLB}(\theta)$ and it is analytically given by [53]:

$$\text{CRLB}(\theta) = \frac{1}{I(\theta)}, \tag{19}$$

where $I(\theta)$ is the so-called Fisher information (FI) for the received data defined as:

$$I(\theta) \triangleq -\mathbb{E}\left\{\frac{\partial^2 \mathcal{L}(\mathbf{Y};\theta)}{\partial\theta^2}\right\}. \tag{20}$$

Recalling the expression of the LLF, $\mathcal{L}(\mathbf{Y};\theta)$, established in (18) and denoting:

$$\gamma_{k,2q}(\theta) \triangleq -\mathbb{E}\Big\{\partial^2 \ln\big(H_{k,2q}\big(u_k(\theta)\big)\big)/\partial\theta^2\Big\},$$
$$\gamma_{k,2q-1}(\theta) \triangleq -\mathbb{E}\Big\{\partial^2 \ln\big(H_{k,2q-1}\big(v_k(\theta)\big)\big)/\partial\theta^2\Big\},$$

it is easy to show the following result:

$$I(\theta) = \sum_{k=0}^{K-1} \big[\gamma_{k,2q}(\theta) + \gamma_{k,2q-1}(\theta)\big]. \tag{21}$$

Clearly, finding the expressions of $\gamma_{k,2q}(\theta)$ and $\gamma_{k,2q-1}(\theta)$ (i.e., carrying out the involved expectations analytically) requires explicit expressions for the pdfs of $u_k(\theta)$ and $v_k(\theta)$. To that end, we show in Appendix A the following result:

Lemma 1: $u_k(\theta)$ and $v_k(\theta)$ are two *independent* RVs whose distributions are given by:

$$p\big[u_k(\theta)\big] = \frac{2\beta_{k,2q}}{\sqrt{\pi\sigma^2}} H_{k,2q}\big(u_k(\theta)\big)\, e^{-\frac{u_k(\theta)^2}{\sigma^2}}, \tag{22}$$

$$p\big[v_k(\theta)\big] = \frac{2\beta_{k,2q-1}}{\sqrt{\pi\sigma^2}} H_{k,2q-1}\big(v_k(\theta)\big)\, e^{-\frac{v_k(\theta)^2}{\sigma^2}}, \tag{23}$$

with

$$\beta_{k,2q} = \frac{1}{2^q}\prod_{l=1}^{q}\frac{1}{\cosh\big(L_{2l}(k)/2\big)} \quad \text{and}$$
$$\beta_{k,2q-1} = \frac{1}{2^q}\prod_{l=1}^{q}\frac{1}{\cosh\big(L_{2l-1}(k)/2\big)}. \tag{24}$$

Proof: See Appendix A.

Due to space limitations, we will hereafter detail the derivation of $\gamma_{k,2q}(\theta)$ only since $\gamma_{k,2q-1}(\theta)$ can be easily obtained in the very same way due to the apparent symmetries between the pdfs of $u_k(\theta)$ and $v_k(\theta)$ [cf. (22) and (23)]. Denoting the first and second derivatives of $H_{k,2q}(x)$ as $H'_{k,2q}(x)$ and $H''_{k,2q}(x)$, respectively, it can be shown that:

$$\frac{\partial^2}{\partial\theta^2}\ln\big(H_{k,2q}(u_k(\theta))\big) = \dot{u}_k^2(\theta)\left[\frac{H''_{k,2q}(u_k(\theta))}{H_{k,2q}(u_k(\theta))} - \frac{H'^2_{k,2q}(u_k(\theta))}{H^2_{k,2q}(u_k(\theta))}\right] + \ddot{u}_k(\theta)\frac{H'_{k,2q}(u_k(\theta))}{H_{k,2q}(u_k(\theta))},$$

where $\dot{u}_k(\theta) = \partial u_k(\theta)/\partial\theta$ and $\ddot{u}_k(\theta) = \partial^2 u_k(\theta)/\partial\theta^2$. Taking the expectation of the above expression and using the fact that $u_k(\theta)$ and $\dot{u}_k(\theta)$ are independent (see Appendix B), we obtain:

$$\gamma_{k,2q}(\theta) = \mathbb{E}\left\{\dot{u}_k^2(\theta)\right\}\left[\mathbb{E}\left\{\frac{H'^2_{k,2q}(u_k(\theta))}{H^2_{k,2q}(u_k(\theta))}\right\} - \mathbb{E}\left\{\frac{H''_{k,2q}(u_k(\theta))}{H_{k,2q}(u_k(\theta))}\right\}\right] - \mathbb{E}\left\{\ddot{u}_k(\theta)\frac{H'_{k,2q}(u_k(\theta))}{H_{k,2q}(u_k(\theta))}\right\}. \tag{25}$$

In the sequel, we will derive the analytical expressions of the four expectations involved in (25), separately. For more convenience, we define beforehand the following two quantities (for $s = 2q$ or $2q-1$) that will appear repeatedly in the obtained results:

$$\omega_{k,s} \triangleq 2\, d_q^2\, \beta_{k,s}\cosh\big(\frac{L_s(k)}{2}\big)\sum\nolimits_{i=1}^{2^{q-1}} \eta_{k,s}(i)(2i-1)^2, \tag{26}$$

$$\alpha_{k,s} \triangleq 2\, d_q\, \beta_{k,s}\sinh\big(\frac{L_s(k)}{2}\big)\sum\nolimits_{i=1}^{2^{q-1}} \eta_{k,s}(i)(2i-1). \tag{27}$$

1) Derivation of $\mathbb{E}\{\dot{u}_k^2(\theta)\}$: Recalling the definition of $u_k(\theta)$ in (17) and using (4), it follows that:

$$\dot{u}_k(\theta) = \frac{S}{\sqrt{N_a}}\Re\left\{\dot{\mathbf{a}}(\theta)^H\mathbf{a}(\theta)x(k)\right\} + \frac{1}{\sqrt{N_a}}\Re\left\{\dot{\mathbf{a}}(\theta)^H\mathbf{w}(k)\right\}, \tag{28}$$

where $\dot{\mathbf{a}}(\theta) = \partial\mathbf{a}(\theta)/\partial\theta$. Then, using the trivial identity $\Re\{z\} = \frac{1}{2}(z+z^*)$ with $z = \dot{\mathbf{a}}(\theta)^H\mathbf{a}(\theta)x(k)$ and the fact that $\dot{\mathbf{a}}(\theta)^H\mathbf{a}(\theta) + \mathbf{a}(\theta)^H\dot{\mathbf{a}}(\theta) = 0$, derived from the identity $\|\mathbf{a}(\theta)\|^2 = N_a$, we obtain:

$$\dot{u}_k(\theta) = \frac{jS}{\sqrt{N_a}}\dot{\mathbf{a}}(\theta)^H\mathbf{a}(\theta)\Im\{x(k)\} + \frac{1}{\sqrt{N_a}}\Re\left\{\dot{\mathbf{a}}(\theta)^H\mathbf{w}(k)\right\}. \tag{29}$$

Owing to the independence between the transmitted symbol, $x(k)$, and the noise vector, $\mathbf{w}(k)$, and noticing that $\dot{\mathbf{a}}(\theta)^H\mathbf{a}(\theta)$ is a pure imaginary number, it can be shown that:

$$\mathbb{E}\left\{\dot{u}_k^2(\theta)\right\} = \frac{S^2}{N_a}\left|\dot{\mathbf{a}}(\theta)^H\mathbf{a}(\theta)\right|^2\mathbb{E}\left\{\Im\{x(k)\}^2\right\} + \frac{1}{N_a}\mathbb{E}\left\{\Re\left\{\dot{\mathbf{a}}(\theta)^H\mathbf{w}(k)\right\}^2\right\}. \tag{30}$$

After tedious algebraic manipulations, we show in Appendix B the following two identities:

$$\mathbb{E}\left\{\Re\left\{\dot{\mathbf{a}}(\theta)^H\mathbf{w}(k)\right\}^2\right\} = \frac{\sigma^2}{2}\left|\left|\dot{\mathbf{a}}(\theta)\right|\right|^2, \tag{31}$$

and

$$\mathbb{E}\left\{\Im\{x(k)\}^2\right\} = \omega_{k,2q-1}. \tag{32}$$

Thus, the analytical expression of $\mathbb{E}\{\dot{u}_k(\theta)^2\}$ is obtained as follows:

$$\mathbb{E}\left\{\dot{u}_k(\theta)^2\right\} = \frac{\sigma^2}{2N_a}\left\|\dot{\mathbf{a}}(\theta)\right\|^2 + \frac{S^2\omega_{k,2q-1}}{2N_a}\left|\mathbf{a}^H(\theta)\dot{\mathbf{a}}(\theta)\right|^2. \tag{33}$$

2) Derivation of $\mathbb{E}\{H'^2_{k,2q}(u_k(\theta))/H^2_{k,2q}(u_k(\theta))\}$: This term is nothing but the expected value of a known transformation of the RV $u_k(\theta)$ whose distribution was already established in (22). Therefore, it can be evaluated analytically by integrating over $p[u_k(\theta)]$ to obtain:

$$\begin{aligned}\mathbb{E}\left\{\frac{H'^2_{k,2q}(u_k(\theta))}{H^2_{k,2q}(u_k(\theta))}\right\} &= \int_{\mathbb{R}}\frac{H'^2_{k,2q}(u_k(\theta))}{H^2_{k,2q}(u_k(\theta))}p[u_k(\theta)]\,\mathrm{d}u_k(\theta)\\ &= \frac{2\beta_{k,2q}}{\sqrt{\pi\sigma^2}}\int_{\mathbb{R}}\frac{H'^2_{k,2q}(u_k(\theta))}{H_{k,2q}(u_k(\theta))}e^{\frac{-u_k^2(\theta)}{\sigma^2}}\mathrm{d}u_k(\theta),\end{aligned} \tag{34}$$

where $H'_{k,2q}(x) \triangleq \partial H_{k,2q}(x)/\partial x$ is given by:

$$\begin{aligned}H'_{k,2q}(x) = \frac{2S\sqrt{N_a}d_q}{\sigma^2}\sum_{i=1}^{2^{q-1}}\eta_{k,2q}(i)(2i-1)^2e^{-\rho N_a d_q^2[2i-1]^2}\\ \times\sinh\left(\frac{2S[2i-1]\sqrt{N_a}d_q}{\sigma^2}x+\frac{L_{2q}(k)}{2}\right).\end{aligned} \tag{35}$$

Injecting (35) back into (34) and using the variable substitution, $t = 2u_k(\theta)/\sigma$, it follows that:

$$\mathbb{E}\left\{\frac{H'^2_{k,2q}(u_k(\theta))}{H^2_{k,2q}(u_k(\theta))}\right\} = \frac{4N_a}{\sigma^2}\rho\,\nu_{k,2q}(\rho), \tag{36}$$

where $\nu_{k,2q}(.)$ is given by:

$$\nu_{k,2q}(\rho) = \frac{d_q^2\beta_{k,2q}}{\sqrt{\pi}}\int_{-\infty}^{+\infty}\frac{f^2_{k,2q}(t,\rho)}{g_{k,2q}(t,\rho)}e^{-\frac{t^2}{4}}\mathrm{d}t, \tag{37}$$

in which $f_{k,s}(t,\rho)$ and $g_{k,s}(t,\rho)$ (for $s = 2q$ and $2q-1$) are expressed as:

$$\begin{aligned}f_{k,s}(t,\rho) = \sum\nolimits_{i=1}^{2^{q-1}}(2i-1)\eta_{k,s}(i)\;e^{-(2i-1)^2d_q^2\rho N_a}\\ \times\sinh\left(\sqrt{2\rho N_a}[2i-1]d_qt+\frac{L_s(k)}{2}\right),\end{aligned} \tag{38}$$

$$\begin{aligned}g_{k,s}(t,\rho) = \sum\nolimits_{i=1}^{2^{q-1}}\eta_{k,s}(i)e^{-(2i-1)^2d_q^2\rho N_a}\\ \times\cosh\left(\sqrt{2\rho N_a}[2i-1]d_qt+\frac{L_s(k)}{2}\right).\end{aligned} \tag{39}$$

3) Derivation of $\mathbb{E}\{H''_{k,2q}(u_k(\theta))/H_{k,2q}(u_k(\theta))\}$: As done in (34), this expectation is found explicitly by integrating over the pdf of $u_k(\theta)$ yielding:

$$\mathbb{E}\left\{\frac{H''_{k,2q}(u_k(\theta))}{H_{k,2q}(u_k(\theta))}\right\} = \int_{\mathbb{R}}\frac{H''_{k,2q}(u_k(\theta))}{H_{k,2q}(u_k(\theta))}p[u_k(\theta)]\mathrm{d}u_k(\theta). \tag{40}$$

Replacing $p[u_k(\theta)]$ with its expression established in (22), it immediately follows that:

$$\mathbb{E}\left\{\frac{H''_{k,2q}(u_k(\theta))}{H_{k,2q}(u_k(\theta))}\right\} = \frac{2\beta_{k,2q}}{\sqrt{\pi\sigma^2}}\int_{\mathbb{R}}H''_{k,2q}\big(u_k(\theta)\big)e^{\frac{-u_k^2(\theta)}{\sigma^2}}\mathrm{d}u_k(\theta), \tag{41}$$

where $H''_{k,2q}(x) \triangleq \partial^2\ H_{k,2q}(x)/\partial x^2$ is given by:

$$\begin{aligned}H''_{k,2q}(x) = \frac{4S^2d_q^2N_a}{\sigma^4}\sum_{i=1}^{2^{q-1}}\eta_{k,2q}(i)(2i-1)^2e^{-\rho N_a d_q^2[2i-1]^2}\\ \times\cosh\left(\frac{2S[2i-1]\sqrt{N_a}d_q}{\sigma^2}x+\frac{L_{2q}(k)}{2}\right).\end{aligned}$$

After expanding the expression of $H''_{k,2q}(x)$ using the identity $\cosh(x+y) = \cosh(x)\cosh(y)+\sinh(x)\sinh(y)$, plugging the result back into (41), and exploiting the fact that $\sinh(x)e^{-\frac{x^2}{2}}$ is an odd function (i.e., its integral is identically zero), we obtain:

$$\begin{aligned}&\mathbb{E}\left\{\frac{H''_{k,2q}\big(u(k)\big)}{H_{k,2q}\big(u_k(\tau)\big)}\right\}\\ &= \frac{8\beta_{k,2q}S^2N_ad_q^2}{(\sqrt{\pi\sigma^2})\sigma^4}\cosh\left(\frac{L_{2q}(k)}{2}\right)\\ &\quad\times\sum_{i=1}^{2^{q-1}}(2i-1)^2\eta_{k,2q}(i)e^{-\rho N_a(2i-1)^2d_q}\\ &\quad\times\int_{\mathbb{R}}\cosh\left(\frac{2S[2i-1]\sqrt{N_a}d_q}{\sigma^2}u_k(\theta)\right)e^{\frac{-u_k^2(\theta)}{\sigma^2}}\mathrm{d}u_k(\theta).\end{aligned} \tag{42}$$

Moreover, we show via "integration by parts" for any $a > 0$ and $b \in \mathbb{R}$ that:

$$\int_0^{+\infty}\cosh(bx)e^{-ax^2}dx = \frac{1}{2}\sqrt{\frac{\pi}{a}}e^{\frac{b^2}{4a}}. \tag{43}$$

With some easy identifications, this result is used in (42) to yield the an analytical expression for the expectation in (41) as follows:

$$\mathbb{E}\left\{\frac{H''_{k,2q}(u_k(\theta))}{H_{k,2q}(u_k(\theta))}\right\} = \frac{4N_a}{\sigma^2}\rho\omega_{k,2q}. \tag{44}$$

4) Derivation of $\mathbb{E}\{\ddot{u}_k(\theta)H'_{k,2q}(u_k(\theta))/H_{k,2q}(u_k(\theta))\}$: To find this expectation, we use a standard probability approach in which we first find the expectation with respect to $\ddot{u}_k(\theta)$ conditioned on $u_k(\theta)$ and then average the result with respect to $u_k(\theta)$. By doing so, we obtain:

$$\begin{aligned}&\mathbb{E}\left\{\ddot{u}_k(\theta)\frac{H'_{k,2q}(u_k(\theta))}{H_{k,2q}(u_k(\theta))}\right\}\\ &\quad= \mathbb{E}_{u_k}\left\{\mathbb{E}_{\ddot{u}_k|u_k}\left\{\ddot{u}_k(\theta)\middle|u_k(\theta)\right\}\frac{H'_{k,2q}(u_k(\theta))}{H_{k,2q}(u_k(\theta))}\right\}.\end{aligned} \tag{45}$$

In Appendix C, we express $\ddot{u}_k(\theta)$ in terms of $u_k(\theta)$ and $v_k(\theta)$ as follows:

$$\ddot{u}_k(\theta) = -\frac{1}{N_a}\left\|\dot{\mathbf{a}}(\theta)\right\|u_k(\theta) - \frac{1}{N_a}\Im\left\{\mathbf{a}(\theta)^H\ddot{\mathbf{a}}(\theta)\right\}v_k(\theta) + z_k(\theta), \tag{46}$$

where $z_k(\theta)$ is another zero-mean RV that is also independent from $u_k(\theta)$. Therefore, using the independence of $u_k(\theta)$ and $v_k(\theta)$ [already established in LEMMA 1], it follows that:

$$\mathbb{E}\left\{\ddot{u}_k(\theta)\middle|u_k(\theta)\right\} = -\frac{1}{N_a}\left\|\dot{\mathbf{a}}(\theta)\right\| u_k(\theta) - \frac{1}{N_a}\Im\left\{\mathbf{a}(\theta)^H\ddot{\mathbf{a}}(\theta)\right\}\mathbb{E}\{v_k(\theta)\}.$$

The expected value of $v_k(\theta)$ can be easily obtained by integrating over its distribution already established in (23) leading to $\mathbb{E}\{v_k(\theta)\} = S\sqrt{N_a}\alpha_{k,2q-1}$. After using these results in (45) and then finding the outer expectation by integration over $p[u_k(\theta)]$, we obtain:

$$\mathbb{E}\left\{\ddot{u}(k)\frac{H'_{k,2q}(u_k(\theta))}{H_{k,2q}(u_k(\theta))}\right\} = -2\rho\left\|\dot{\mathbf{a}}(\theta)\right\|^2\omega_{k,2q} - 2\rho\Im\left\{\mathbf{a}(\theta)^H\ddot{\mathbf{a}}(\theta)\right\}\alpha_{k,2q}\alpha_{k,2q-1}. \tag{47}$$

Plugging the four expectations established in (33), (36), (44), and (47) back into (25), it follows that:

$$\gamma_{k,2q}(\theta) = 2\rho^2\omega_{k,2q-1}\left|\mathbf{a}^H(\theta)\dot{\mathbf{a}}(\theta)\right|^2\left[\nu_{k,2q}(\rho) - \omega_{k,2q}\right] + 2\rho\left\|\dot{\mathbf{a}}(\theta)\right\|^2\nu_{k,2q}(\rho) + 2\rho\Im\left\{\mathbf{a}(\theta)^H\ddot{\mathbf{a}}(\theta)\right\}\alpha_{k,2q}\alpha_{k,2q-1}. \tag{48}$$

As mentioned previously, due to the apparent symmetries between the distributions of $u_k(\theta)$ and $v_k(\theta)$, the analytical expression of $\gamma_{k,2q-1}(\theta)$ can be directly deduced from $\gamma_{k,2q}(\theta)$ by easy identifications as follows:

$$\gamma_{k,2q-1}(\theta) = 2\rho^2\omega_{k,2q}\left|\mathbf{a}^H(\theta)\dot{\mathbf{a}}(\theta)\right|^2\left[\nu_{k,2q-1}(\rho) - \omega_{k,2q-1}\right] + 2\rho\left\|\dot{\mathbf{a}}(\theta)\right\|^2\nu_{k,2q-1}(\rho) - 2\rho\Im\left\{\mathbf{a}(\theta)^H\ddot{\mathbf{a}}(\theta)\right\}\alpha_{k,2q}\ \alpha_{k,2q-1}. \tag{49}$$

The DOA CA CRLBs are then simply obtained as follows:

$$\mathrm{CRLB}(\theta) = \frac{1}{\sum_{k=0}^{K-1}\left[\gamma_{k,2q}(\theta) + \gamma_{k,2q-1}(\theta)\right]}. \tag{50}$$

It is worth mentioning here that we have only exploited the fact that the constellation is Gray-coded so far, and that the specific turbo-code structure and setup was not truly used during all previous derivations. Therefore, the new CA CRLB expression in (50) is actually valid for any coded system in general. Yet, the CRLB was explicitly expressed in terms of the *a priori* LLRs of the coded bits through the coefficients $\omega_{k,s}$, $\nu_{k,s}$, and $\alpha_{k,s}$ (for $s = 2q$ and $s = 2q-1$) that are involved in the expressions of $\gamma_{k,2q}(\theta)$ and $\gamma_{k,2q-1}(\theta)$. However, these unknown LLRs and, therefore, the CRLB cannot be evaluated for any code. Yet, we will soon explain (cf. Section IV-B) how these unknown LLRs can be obtained from the output of the SISO decoders in a turbo-coded system.

Interestingly enough, the established analytical CRLBs are also valid for any antenna array geometry. Yet, uniform linear arrays (ULAs) and uniform circular arrays (UCAs) remain by far the most studied cases in the open literature. For these two popular configurations, by assuming that the distance between the antenna elements is half the wavelength, the corresponding transformation of the DOA parameter in (5) is given by:

$$\varphi_i^{\mathrm{ULA}}(\theta) = (i-1)\sin(\theta),$$
$$\varphi_i^{\mathrm{UCA}}(\theta) = \frac{\cos\left(\theta - 2[i-1]\pi/N_a\right)}{2\sin(\pi/N_a)}. \tag{51}$$

Therefore, it can be easily shown that the geometrical factors $|\mathbf{a}(\theta)^H\dot{\mathbf{a}}(\theta)|^2$ and $\left\|\dot{\mathbf{a}}(\theta)\right\|^2$ involved in $\gamma_{k,2q}(\theta)$ and $\gamma_{k,2q-1}(\theta)$ are explicitly expressed as in Table I. The contribution of the other geometrical factor, $\Im\left\{\mathbf{a}(\theta)^H\ddot{\mathbf{a}}(\theta)\right\}$, cancels out by adding $\gamma_{k,2q}(\theta)$ and $\gamma_{k,2q-1}(\theta)$ in (50) and, therefore, it is not shown in Table I.

TABLE I
GEOMETRICAL FACTORS FOR ULA AND UCA CONFIGURATIONS

	$\mathbb{ULA}$	$\mathbb{UCA}$
$\left\|\dot{\mathbf{a}}(\theta)\right\|^2$	$\pi^2\frac{N_a(N_a-1)(2N_a-1)}{6}\cos^2(\theta)$	$\frac{N_a\pi^2}{8\sin^2(\pi/N_a)}$
$\left\|\mathbf{a}(\theta)^H\dot{\mathbf{a}}(\theta)\right\|^2$	$\pi^2\cos^2(\theta)\left(\sum_{k=1}^{N_a-1}k\right)^2$	0

We emphasize here the fact that the CRLB depends on the DOA parameter only through these purely geometrical factors which are: i) even functions of θ for ULAs thereby reflecting their symmetry with respect to the broadside axis, and ii) independent of θ for UCAs owing to their circular symmetry.

B. Evaluation of the Analytical CA CRLBs in Turbo-Coded Transmissions

As already mentioned, in order to compute and plot the new CA CRLBs, one needs to evaluate the coefficients $\omega_{k,s}$, $\nu_{k,s}$, and $\alpha_{k,s}$ for $s = 2q$ and $s = 2q-1$. These coefficients are, however, functions of the *a priori* LLRs, $L_l(k)$. In the sequel, we briefly explain how these LLRs can be obtained from the output of the SISO decoders after convergence of the BCJR algorithm. First, a receive beamformer steered toward the true DOA, θ, returns the information-bearing sequence:

$$\mathbf{y}(\theta) = \left[y_0(\theta), y_1(\theta), \ldots, y_{K-1}(\theta)\right]^T, \tag{52}$$

where

$$y_k(\theta) = \frac{1}{\sqrt{N_a}}\mathbf{a}^H(\theta)\mathbf{y}(k). \tag{53}$$

Then, the demapper extracts the so-called *bit likelihoods*:

$$\Lambda_l(k) \triangleq \ln\left(\frac{p\left[\mathbf{y}(\theta)\middle|b_l^k = 1\right]}{p\left[\mathbf{y}(\theta)\middle|b_l^k = 0\right]}\right), \tag{54}$$

for all the coded bits and feed them as inputs to the turbo decoder. In practice, $\Lambda_l(k)$, are explicitly expressed as well

using the assumption of independent symbols as follows[2] [60]:

$$\Lambda_l(k) = \ln\left(\frac{\sum_{c_m \in \mathcal{C}_l^{(1)}} \exp\left\{-\frac{1}{\sigma^2}\left|y_k(\theta) - S\sqrt{N_a}c_m\right|^2\right\}}{\sum_{c_m \in \mathcal{C}_l^{(0)}} \exp\left\{-\frac{1}{\sigma^2}\left|y_k(\theta) - S\sqrt{N_a}c_m\right|^2\right\}}\right), \tag{55}$$

where $\mathcal{C}_l^{(1)}$ (resp. $\mathcal{C}_l^{(0)}$) is the subset of constellation points whose l^{th} mapped bit is equal to 1 (resp. 0). In practice, (55) is easily computed for all l and k using the so-called max-log-MAP approximation (cf. [60]. pp 24-26] for more details). Then, by exchanging the so-called *extrinsic* information between the two SISO decoders, the *a posteriori* LLRs of the coded bits:

$$\Upsilon_l(k) = \ln\left(\frac{Pr\left[b_l^k = 1\middle|\mathbf{y}(\theta)\right]}{Pr\left[b_l^k = 0\middle|\mathbf{y}(\theta)\right]}\right), \tag{56}$$

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

$$Pr\left[b_l^k = 1\middle|\mathbf{y}(\theta)\right] = \frac{p\left[\mathbf{y}(\theta)\middle|b_l^k = 1\right]Pr\left[b_l^k = 1\right]}{p[\mathbf{y}(\theta)]}, \tag{57}$$

and

$$Pr\left[b_l^k = 0\middle|\mathbf{y}(\theta)\right] = \frac{p\left[\mathbf{y}(\theta)\middle|b_l^k = 0\right]Pr\left[b_l^k = 0\right]}{p[\mathbf{y}(\theta)]}. \tag{58}$$

Therefore, by taking the ratio of (57) and (58) and applying the natural logarithm, it immediately follows that:

$$L_l(k) = \Upsilon_l(k) - \Lambda_l(k) \;\approx\; \Upsilon_l^{(R)}(k) - \Lambda_l(k), \tag{59}$$

meaning that the required *a priori* LLRs of the coded bits can be easily obtained from both their steady-state *a posteriori* LLRs and $\Lambda_l(k)$ already computed by the *soft* demapper prior to data decoding. The CA CRLBs can be readily computed after using $L_l(k)$ to evaluate $\omega_{k,s}$, $\alpha_{k,s}$, and $\nu_{k,s}(\rho)$ for $s = 2q$ and $2q-1$.

C. Empirical CA CRLBs

For validation purposes, we develop another approach that evaluates the considered bounds via exhaustive Monte-Carlo simulations. To that end and as done in [56] in the context of time and frequency synchronization, we also use the following alternative expression for the LLF:

$$\mathcal{L}(\mathbf{Y};\theta) = \ln\Big(\sum\nolimits_{l=1}^{M^K} Pr[\mathbf{x} = \mathbf{x}_l]\, p[\mathbf{Y}|\mathbf{x} = \mathbf{x}_l;\theta]\Big), \tag{60}$$

where $\{\mathbf{x}_l\}_{l=1}^{M^K}$ stand for all the possible values that the whole transmitted symbol sequence $\mathbf{x} = [x(0), x(1), \ldots, x(K-1)]^T$ can take. Then, noticing that:

$$\frac{\partial p[\mathbf{Y}|\mathbf{x} = \mathbf{x}_l;\theta]}{\partial\theta} = p[\mathbf{Y}|\mathbf{x} = \mathbf{x}_l;\theta]\frac{\partial \ln\left(p[\mathbf{Y}|\mathbf{x} = \mathbf{x}_l;\theta]\right)}{\partial\theta}, \tag{61}$$

it follows that the first derivative of (60) with respect to θ is obtained as:

$$\frac{\partial\mathcal{L}(\mathbf{Y};\theta)}{\partial\theta} = \sum_{l=1}^{M^K}\frac{Pr[\mathbf{x} = \mathbf{x}_l]\, p[\mathbf{Y}|\mathbf{x} = \mathbf{x}_l;\theta]}{p[\mathbf{Y};\theta]} \times\frac{\partial \ln\left(p[\mathbf{Y}|\mathbf{x} = \mathbf{x}_l;\theta]\right)}{\partial\theta}.$$

Besides, similar to [56] and using the Bayes' formula, it can be shown that:

$$\begin{aligned}\frac{\partial\mathcal{L}(\mathbf{Y};\theta)}{\partial\theta} &= E_{\mathbf{x}|\mathbf{Y}}\left\{\frac{\partial \ln(p[\mathbf{Y}|\mathbf{x};\theta])}{\partial\theta}\right\}\\ &= \sum_{k=0}^{K-1} E_{x(k)|\mathbf{Y}}\Big\{\frac{2S}{\sigma^2}\Re\left\{\dot{\mathbf{a}}(\theta)^H\mathbf{y}(k)x(k)^*\right\}\Big\}.\end{aligned} \tag{62}$$

The *a posteriori* probabilities, $Pr[x(k) = c_m|\mathbf{Y}]$, obtained from the output of the SISO decoders, are used to compute the elementary conditional expectations in (62) as follows:

$$\begin{aligned}&\mathbb{E}_{x(k)|\mathbf{Y}}\Big\{\frac{2S}{\sigma^2}\Re\left\{\dot{\mathbf{a}}(\theta)^H\mathbf{y}(k)x(k)^*\right\}\Big\}\\ &\quad= \sum_{c_m\in\mathcal{C}} Pr[x(k) = c_m|\mathbf{Y}]\Big(\frac{2S}{\sigma^2}\Re\left\{\dot{\mathbf{a}}(\theta)^H\mathbf{y}(k)c_m^*\right\}\Big).\end{aligned} \tag{63}$$

Then, a large number of noisy realizations, $\{\mathbf{Y}_l\}_{l=1}^L$, for $\mathbf{Y}$ are generated and the considered CA CRLBs are evaluated *empirically* as follows (cf. [56] for more details):

$$\text{CRLB} = \left(\frac{1}{L}\sum_{l=1}^{L}\left(\frac{\partial\mathcal{L}(\mathbf{Y}_l;\theta)}{\partial\theta}\right)^2\right)^{-1}. \tag{64}$$

V. CA DOA ML Estimator

As mentioned in Section II, all existing direction finding techniques are completely oblivious to the decoder. In fact, with no *a priori* knowledge about the transmitted symbols, the receiver starts by estimating the DOA directly from the snapshots, $\{\mathbf{y}(k)\}_{k=0}^{K-1}$. Then, a receive beamformer steered toward the acquired NDA DOA estimate, $\widehat{\theta}_{\text{NDA}}$, returns the information-bearing sequence, $\mathbf{y}(\widehat{\theta}_{\text{NDA}})$, that is used to decode the data. As a matter of fact, $\widehat{\theta}_{\text{NDA}}$ can be obtained by applying any of the classical direction finding techniques. Yet, a more accurate *stochastic* ML estimate (MLE) for θ is obtained in this paper as follows:

$$\widehat{\theta}_{\text{ML-NDA}} = \underset{\theta'}{\text{argmax}}\;\mathcal{L}^{(0)}(\theta'), \tag{65}$$

where $\mathcal{L}^{(0)}(.)$ is the NDA LLF obtained directly from its CA counterpart in (18) by setting[3] $L_l(k) = 0$ for all l and k, i.e.:

$$\mathcal{L}^{(0)}(\theta') = \sum_{k=0}^{K-1}\Big[\ln\left(H\left(u_k(\theta')\right)\right) \;+\; \ln\left(H\left(v_k(\theta')\right)\right)\Big], \tag{66}$$

[2]Note here that the analytical CA CRLBs estbalished in this paper are valid for Gray mapping only which is itself optimal for *non-iterative* BICM schemes only. For this reason, no feedback from the decoder is exploited by the demapper as evidenced by (55). For BICM systems with iterative demapping/demodulation (i.e., BICM-ID), however, set-partitioning (SP) labelling is the optimum mapping scheme. Generalization of our derivations (especially the likelihood factorization) to the latter case in conjunction with BICM-ID schemes falls, however, beyond the scope of this paper and will be considered in a future work.

[3]In the NDA case (i.e., before starting data decoding), no *a priori* information about the bits is available at the receiver end, i.e., $P[b_l^k = 0] = P[b_l^k = 1] = 1/2$ and therefore, owing to (1), $L_l(k) = 0$ for all l and k.

in which $H(.)$ is simply given by:

$$H(x) = \sum_{i=1}^{2^{q-1}} e^{-\rho N_a d_q^2 [2i-1]^2} \cosh\left(\frac{2S[2i-1]\sqrt{N_a} d_q}{\sigma^2} x\right).$$

In (66), $u_k(\theta')$ and $v_k(\theta')$ are the real and imaginary parts of:

$$y_k(\theta') = \frac{1}{\sqrt{N_a}} \mathbf{a}^H(\theta') \mathbf{y}(k), \tag{67}$$

which is the output of a receive beamformer steered toward any candidate direction, θ', that the true unknown DOA, θ, can possibly take. Clearly, the receive beamformer maximizes the antenna gain only when $\theta' = \theta$ leading indeed to the NDA MLE in (65). The beamforming output sequence:

$$\mathbf{y}(\widehat{\theta}_{\text{ML-NDA}}) = \left[y_0(\widehat{\theta}_{\text{ML-NDA}}), y_1(\widehat{\theta}_{\text{ML-NDA}}), \ldots, y_{K-1}(\widehat{\theta}_{\text{ML-NDA}})\right]^T,$$

corresponding to $\widehat{\theta}_{\text{ML-NDA}}$ can be used to decode the data once for all as explained in subsection IV-B. By doing so, however, the system is prone to severe decoding errors if $\widehat{\theta}_{\text{ML-NDA}}$ is not sufficiently accurate due to harsh SNR conditions. In the sequel, we will introduce a novel and more elaborate solution wherein the sequence $\mathbf{y}(\widehat{\theta}_{\text{ML-NDA}})$ is used to initiate the decoding process only. The direction finding and the beamforming tasks are then embedded within the turbo iteration loop. To do so, we modify (59) as follows:

$$L_l^{(r)}(k) = \Upsilon_l^{(r)}(k) - \Lambda_l^{(r-1)}(k), \tag{68}$$

in order to obtain the *a priori* LLRs and, hence, a more refined DOA MLE, $\widehat{\theta}_{\text{ML-CA}}^{(r)}$, at each r^{th} turbo iteration. It is worth mentioning here that the bit likelihoods, $\Lambda_l^{(r-1)}(k)$, in (68) are extracted[4] from the beamforming output sequence, $\mathbf{y}(\widehat{\theta}_{\text{ML-CA}}^{(r-1)})$, obtained during the previous turbo iteration. The DOA MLE, $\widehat{\theta}_{\text{ML-CA}}^{(r)}$, pertaining to the r^{th} turbo iteration is obtained as follows:

$$\widehat{\theta}_{\text{ML-CA}}^{(r)} = \underset{\theta'}{\text{argmax}}\ \mathcal{L}^{(r)}(\theta'), \tag{69}$$

where $\mathcal{L}^{(r)}(\theta')$ is the CA LLF in (18) but evaluated using $L_l^{(r)}(k)$ instead of $L_l(k)$, i.e.:

$$\mathcal{L}^{(r)}(\theta') = \sum_{k=1}^{K} \ln\left(H_{k,2q}^{(r)}\left(u_k(\theta')\right)\right) + \ln\left(H_{k,2q-1}^{(r)}\left(v_k(\theta')\right)\right), \tag{70}$$

in which $H_{k,s}^{(r)}(.)$ is given for $s = 2q$ and $2q-1$ by:

$$H_{k,s}^{(r)}(x) = \sum_{i=1}^{2^{q-1}} \eta_{k,s}^{(r)}(i) e^{-\rho N_a [2i-1]^2 d_q^2} \times \cosh\left(\frac{2S[2i-1]\sqrt{N_a} d_q}{\sigma^2} x + \frac{L_s^{(r)}(k)}{2}\right).$$

Here, $\eta_{k,2q}^{(r)}(i)$ and $\eta_{k,2q-1}^{(r)}(i)$ are also obtained by using $L_l^{(r)}(k)$ instead of $L_l(k)$ in (14). A missing detail that needs to be addressed, though, is how the objective CA LLFs in (70) are maximized at each r^{th} turbo iteration. Actually, since they were derived analytically in this paper, these CA LLFs can be easily maximized using any of the popular iterative techniques such as the well-known Newton-Raphson algorithm:

$$\widehat{\theta}_i^{(r)} = \widehat{\theta}_{i-1}^{(r)} - \left[\left(\frac{\partial^2 \mathcal{L}^{(r)}(\theta)}{\partial \theta^2}\right)^{-1} \frac{\partial \mathcal{L}^{(r)}(\theta)}{\partial \theta}\right]_{\theta = \widehat{\theta}_{i-1}^{(r)}}, \tag{71}$$

in which $\widehat{\theta}_i^{(r)}$ is the DOA estimate pertaining to the i^{th} Newton-Raphson iteration. The algorithm stops once the convergence criterion $|\widehat{\theta}_i^{(r)} - \widehat{\theta}_{i-1}^{(r)}| \leq \epsilon$ is met to produce $\widehat{\theta}_{\text{ML-CA}}^{(r)}$ as the CA DOA MLE during the r^{th} turbo iteration. Note here that ϵ is a predefined threshold that governs the required estimation accuracy.

Note also that the Newton-Raphson algorithm itself is iterative in nature and, therefore, requires a reliable initial guess, $\widehat{\theta}_0^{(r)}$, to ensure its convergence to the global maximum of the underlying objective LLF. At each r^{th} turbo iteration, the algorithm is initialized by $\widehat{\theta}_0^{(r)} = \widehat{\theta}_{\text{ML-CA}}^{(r-1)}$ (i.e., by the DOA MLE pertaining to the previous turbo iteration). At the very first turbo iteration, however, the algorithm is initialized with the NDA MLE, $\widehat{\theta}_{\text{ML-NDA}}$, obtained in (65). The latter is obtained by maximizing $\mathcal{L}^{(0)}(\theta)$ itself via the very same Newton-Raphson algorithm and the corresponding initial guess is obtained by a broad line search over θ.

The new CA DOA direction finding scheme is illustrated in Fig. 1. As seen there, the soft demapper extracts the *bit likelihoods* from the previous beamforming output sequence, $\mathbf{y}(\widehat{\theta}_{i-1}^{(r-1)})$, and feed them as inputs to the SISO decoders which compute the *a posteriori* LLRs, $\Upsilon^{(r)}$, at the current turbo iteration. To be more precise, at each r^{th} turbo iteration, the first SISO decoder returns the *a posteriori* LLRs, $\Upsilon_1^{(r)}$, of both the data and parity bits stemming from the first RSC encoder. Only the *a posteriori* LLRs corresponding to the data bits are extracted and used to find their *extrinsic* information, which are then passed on as inputs to the second SISO decoder. The latter also produces the *a posteriori* LLRs, $\widehat{\Upsilon}_2^{(r)}$, of the very same data bits and the parity bits stemming from the second RSC encoder. Then, $\widehat{\Upsilon}_1^{(r)}$ and $\widehat{\Upsilon}_2^{(r)}$, are punctured in order to obtain the *a posteriori* LLRs, $\widehat{\Upsilon}^{(r)}$, of all the coded bits. Their *a priori* LLRs are then acquired as shown in Fig. 1 and fed to the estimation block in order to find the r^{th} CA DOA MLE, $\widehat{\theta}_{\text{ML-CA}}^{(r)}$.

VI. Simulation Results

In this section, we resort to computer simulations to assess the performance of the new CA ML estimator and to validate the new analytical CA CRLBs. The encoder is composed of two identical RSCs of generator polynomials $(1, 0, 1, 1)$ and $(1, 1, 0, 1)$, each with systematic rate $R_0 = 1/2$. 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 RSCs' memory is fixed to 4. We also consider 16− and 64−QAM as representative examples for square-QAM constellations. The number of antennas is fixed to $N_a = 4$ and all the figures are obtained for $\theta = 45°$.

[4]Note here that $\Lambda_l^{(r-1)}(k)$ are extracted according to (54) where $\mathbf{y}(\theta)$ is being replaced by $\mathbf{y}(\widehat{\theta}_{\text{ML-CA}}^{(r-1)})$.

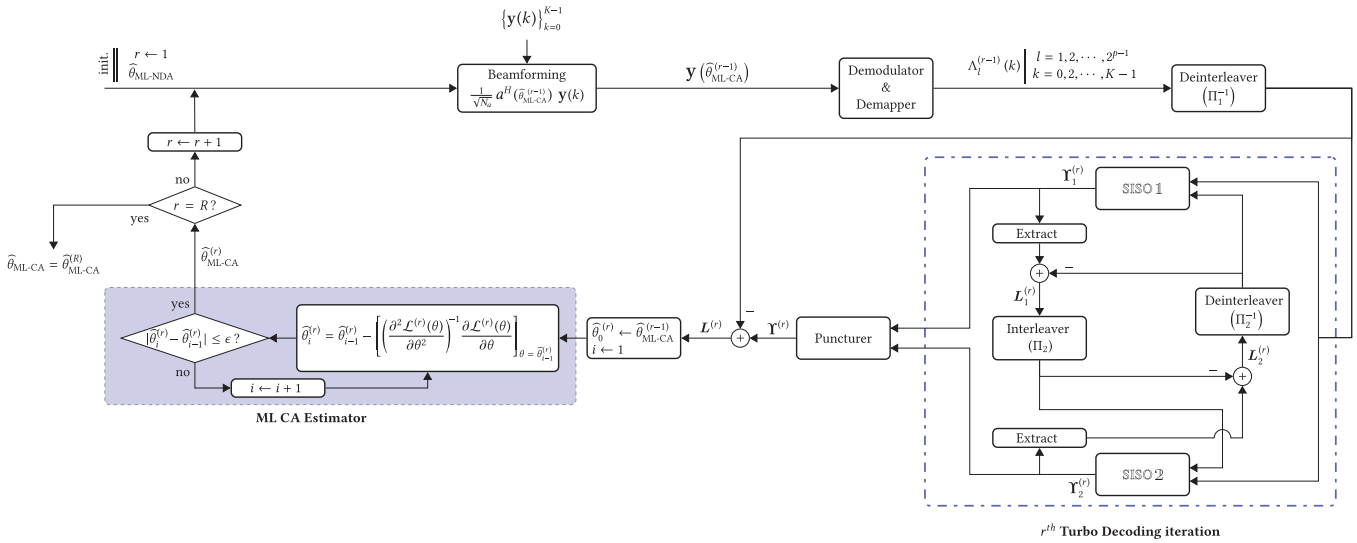


Fig. 1. Architecture of joint CA DOA estimation and turbo decoding.

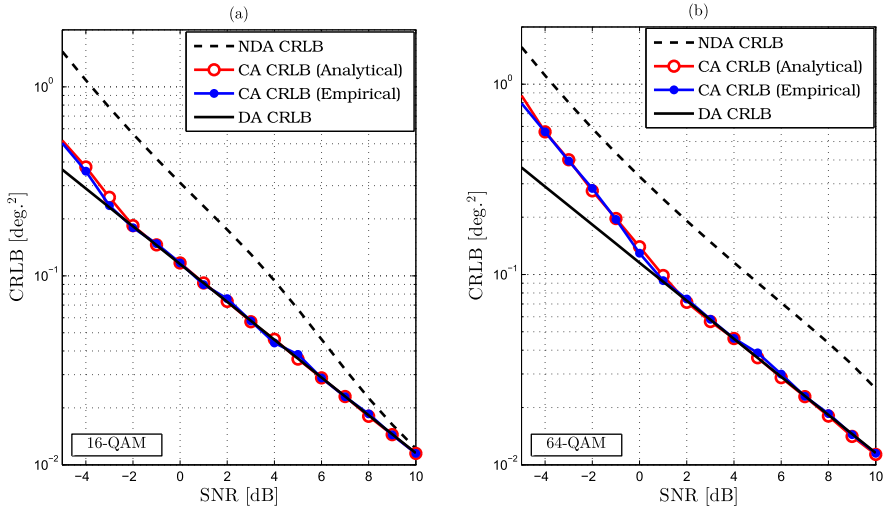


Fig. 2. NDA, DA, and CA (analytical and empirical) CRLBs with $R = 1/3$ and $K = 206$ for: (a) 16-QAM, and (b) 64-QAM.

A. Validation of the New Analytical CA CRLBs

We mention beforehand that the integrand function in (37) takes extremely small values as $|t|$ increases. Therefore, its integral over $(-\infty, +\infty)$ can be accurately approximated by a finite integral over an appropriate support, $[-T, T]$, for which the Riemann integration method can be adequately used. In our simulations, it should be noted that $T = 20$ and a summation step of 1 provided very accurate values for the infinite integral.

In Fig. 2, we plot the *empirical* and *analytical* CA CRLBs together for each modulation order. There, we can see that the analytical CA CRLBs coincide exactly with their empirical counterparts validating thereby our new expressions established in Section III. Besides, as intuitively expected, the CA CRLBs are lower than the NDA CRLBs since the latter bounds correspond to a *completely* blind estimation scenario where no *a priori* information about the transmitted data is exploited at the receiver. This highlights the potential improvements in estimation performance that can be achieved owing to the decoder's assistance. More interestingly, contrarily to the NDA CRLBs, the CA CRLBs decay rapidly and reach the DA bounds starting from very low SNR levels. Therefore, over a wide range of practical SNRs, CA direction finding is theoretically able to achieve the same estimation performance that would be obtained if all the transmitted symbols (or bits) were perfectly known to the receiver.

B. ML Estimator

Figs. 3 and 4 depict the performance of the new CA ML estimator, for 16− and 64−QAM (with two different coding rates each) in terms of the mean square error (MSE). The CA CRLBs are plotted as well to serve as overall benchmarks for the performance of the new estimator under each simulation setup. In order to highlight the performance improvements brought by CA estimation over NDA estimation, from an algorithmic point of view, we also plot in the same figures the MSE of the completely NDA ML estimator, $\widehat{\theta}_{\text{ML-NDA}}$, given

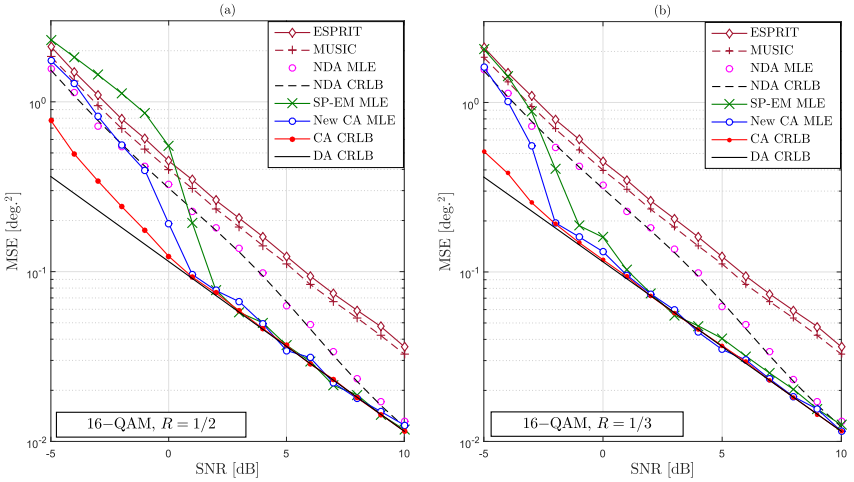


Fig. 3. MSE for the different DOA estimators for 16−QAM, $K = 206$ with: (a) $R = 1/2$, and (b) $R = 1/3$.

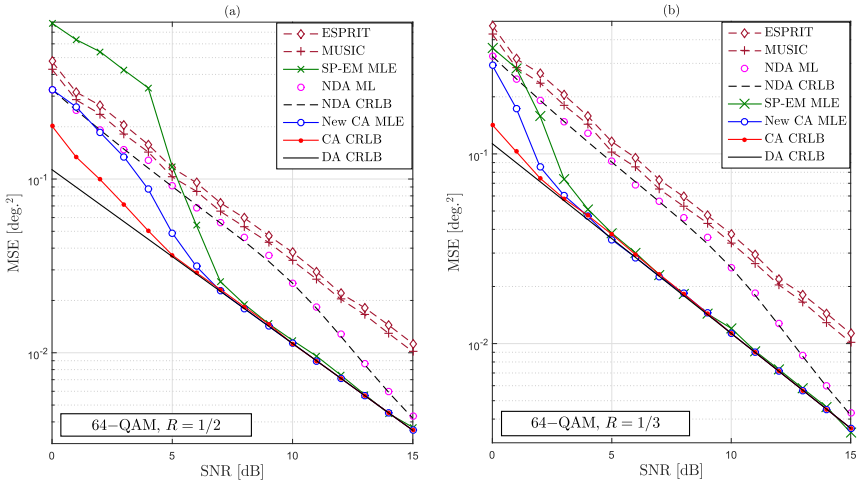


Fig. 4. MSE for the different DOA estimators for 64−QAM, $K = 206$ with: (a) $R = 1/2$, and (b) $R = 1/3$.

by (65). In the same figures, we also compare the new CA DOA MLE to MUSIC, ESPRIT, and SP-EM.

First, it is observed that the new CA ML estimator is able to achieve the performance improvements that were theoretically predicted by the CA CRLBs. Moreover, the fact that the CA ML estimator achieves the CA CRLBs confirms its statistical efficiency in practice. The impact of the coding rate on the estimation performance is also clearly observed from the same figures. In fact, as intuitively expected, the performance of the new CA ML estimator improves, in the low-SNR region, by decreasing the coding rate.[5] This is actually hardly surprising since more error-free bits are detected by the decoder if more redundancy is introduced by the encoder. More specifically, higher redundancy leads to better estimates for the *a priori* LLRs involved in the CA LLF of the system. For the same coding rate, however, the performance of the CA ML estimator (as well as the CA CRLBs) deteriorates by increasing the modulation order at any SNR value. This is a typical behavior that is usually observed in any parameter estimation problem involving linearly-modulated signals (even in the NDA case). Indeed, when the modulation order increases, the inter-symbol distance decreases for normalized-energy constellations. As such, at the same SNR level, noise components have a relatively worse impact on symbol detection and parameter estimation in general.

We also observe from Figs. 3 and 4 that the new CA MLE outperforms the traditional subspace-based approaches (i.e., MUSIC and ESPRIT)[6] over the entire SNR range. This is actually hardly surprising since MUSIC and ESPRIT are completely blind estimators and, as such, their performance is always lower-bounded by the NDA CRLB irrespectively of the number of snapshots and/or the SNR level. From Figs. 3 and 4, we also observe that the proposed CA MLE outperforms SP-EM over the low-SNR range and that this advantage is more prominent at higher-modulation orders and/or lower coding rates. The superiority of the newly proposed CA MLE

[5]Note here that the drastic MSE drop of both SP-EM and the new CA ML algorithm stems from the drastic drop of the bit error rate (BER) with increasing SNR in turbo-coded systems (cf. [10]).

[6]Note here that ESPRIT was implemented using the following configuration: two overlapping subarrays with distance $d_s = 1$ between them. For more details, please see [2].

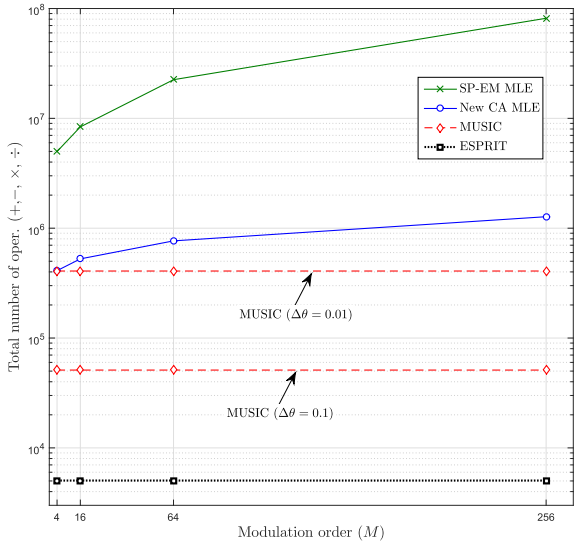


Fig. 5. Complexities of the new CA MLE, SP-EM MLE, ESPRIT ($N_s = 3$, $d_s = 1$), and MUSIC algorithms versus the modulation order M for $K = 206$ and $N_a = 4$.

over SP-EM is, however, much further appreciated when it comes to computational complexity. Indeed, we plot in Fig. 5 the total number of operations (i.e., $+$, $\times$, and $\div$) required by all estimators versus the modulation order. There, we can see that the proposed estimator is about 30 and 65 times computationally less expensive than SP-EM for 64-QAM and 256 QAM, respectively. Being less accurate than the proposed CA ML estimator, the traditional subspace-based algorithms are, however, computationally more attractive. Yet, it should be mentioned that the complexity of MUSIC algorithm is mainly governed by the discretization step $\Delta\theta$ since its spatial spectrum needs to be evaluated at $N = \frac{180}{\Delta\theta}$ candidate values, θ_n, for the true DOA parameter. As such, its complexity increases by decreasing $\Delta\theta$ for the sake of enhanced performance. ESPRIT, however, does not involve any grid search operation and, as such, entails very low computational cost as compared to the proposed CA ML estimator.[7]

VII. Conclusion

In this contribution, a new CA ML DOA estimator was designed for turbo-coded square-QAM transmissions. The direction finding and receive-beamforming tasks were both incorporated within the turbo iteration loop. The exact CRLBs for the underlying estimation problem (i.e., CA CRLBs) were also established analytically. The new estimator reaches the CA CRLBs over a wide range of practical SNRs confirming thereby its statistical efficiency in practice. It also outperforms the *completely* blind (i.e., NDA) ML estimator over the entire SNR range. Most remarkably, with no *a priori* knowledge about the transmitted data, the new CA ML estimator achieves the DA CRLB which corresponds to an ideal scenario where all the transmitted symbols are perfectly known to the receiver. As intuitively expected, its performance further improves at relatively higher coding rates and lower modulation orders. Finally, it is worth mentioning that the results disclosed in this paper can be extended, using equivalent manipulations, to uniform rectangular arrays (URAs) where both the elevation and azimuth angles are estimated. Yet, this extension has its own technical pitfalls such as the challenging pairing problem. Addressing this kind of issues in the most appropriate way falls beyond the scope of this paper and will be considered in a future work.

[7]In practice, opting to either of the methods depends on the specific application at hand and its associated performance/complexity tradeoffs. Indeed, the new CA estimator is more appealing in situations where pointing accuracy is of utmost priority such as in location-aware and 911-emergency services, mmWave or optical wireless (OW) beamforming communication links, etc. ESPRIT, is however, more appropriate when computational complexity is of first concern. It is worth mentioning here that the difference between the complexities of the proposed technique and ESPRIT-type methods will be higher in the multiple-parameter estimation case (e.g., both azimuth and elevation angle estimation with 2D antenna arrays).

Appendix A
Proof of Lemma 1

• *A.1) Independence of $u_k(\theta)$ and $v_k(\theta)$:* We have from (17):

$$u_k(\theta) \triangleq \frac{1}{\sqrt{N_a}}\Re\left\{\mathbf{a}(\theta)^H\mathbf{y}(k)\right\}$$
$$v_k(\theta) \triangleq \frac{1}{\sqrt{N_a}}\Im\left\{\mathbf{a}(\theta)^H\mathbf{y}(k)\right\}. \quad (72)$$

Since $\mathbf{y}(k) = S\,\mathbf{a}(\theta)\,x(k) + \mathbf{w}(k)$, it follows that:

$$u_k(\theta) = S\sqrt{N_a}\,\Re\{x(k)\} + \frac{1}{\sqrt{N_a}}\Re\left\{\mathbf{a}(\theta)^H\mathbf{w}(k)\right\}, \quad (73)$$
$$v_k(\theta) = S\sqrt{N_a}\Im\{x(k)\} + \frac{1}{\sqrt{N_a}}\Im\left\{\mathbf{a}(\theta)^H\mathbf{w}(k)\right\}. \quad (74)$$

Similarly to [33], we take advantage from the independence of $\Re\{x(k)\}$ and $\Im\{x(k)\}$. We verify as well that $x(k)$ and $x(k)^*$ are both independent from $\Re\left\{\mathbf{a}(\theta)^H\mathbf{w}(k)\right\}$ and $\Im\left\{\mathbf{a}(\theta)^H\mathbf{w}(k)\right\}$. Besides, the two RVs $\Re\left\{\mathbf{a}(\theta)^H\mathbf{w}(k)\right\}$ and $\Im\left\{\mathbf{a}(\theta)^H\mathbf{w}(k)\right\}$ are uncorrelated Gaussian RVs and, hence, independent. Therefore, $u_k(\theta)$ and $v_k(\theta)$ are independent.

• *A.2) pdfs of $u_k(\theta)$ and $v_k(\theta)$:* We have from (4):

$$\mathbf{y}(k) = S\mathbf{a}(\theta)x(k) + \mathbf{w}(k). \quad (75)$$

By denoting $y_k(\theta) = \frac{1}{\sqrt{N_a}}\mathbf{a}(\theta)^H\mathbf{y}(k)$, it follows that:

$$y_k(\theta) = S\sqrt{N_a}x(k) + \underbrace{\frac{1}{\sqrt{N_a}}\mathbf{a}(\theta)^H\mathbf{w}(k)}_{\widetilde{w}_k(\theta)},$$

where $\widetilde{w}_k(\theta)$ is a zero-mean Gaussian RV whose variance is given by:

$$\begin{aligned}\mathrm{var}\{\tilde{w}(\theta)\} &= \frac{1}{N_a}\mathbb{E}\left\{\left[\mathbf{a}(\theta)^H\mathbf{w}(k)\right]\left[\mathbf{a}(\theta)^H\mathbf{w}(k)\right]^*\right\}\\ &= \frac{1}{N_a}\mathbf{a}(\theta)^H\mathbb{E}\left\{\mathbf{w}(k)\mathbf{w}(k)^H\right\}\mathbf{a}(\theta) = \sigma^2.\end{aligned}$$

Therefore, the pdf of the RV $y_k(\theta)$ in (76), conditioned on $x(k) = c_m\ \forall\ c_m \in \mathcal{C}$, is given by:

$$p[y_k(\theta)|x(k) = c_m] = \frac{1}{\pi\sigma^2}\exp\left\{-\frac{1}{\sigma^2}\left|y_k(\theta) - S\sqrt{N_a}c_m\right|^2\right\}. \quad (76)$$

By noticing that $y_k(\theta) = u_k(\theta) + jv_k(\theta)$, recalling the expression of $D_k(\theta)$ in (11), and averaging (76) over the constellation alphabet, it follows that:

$$p[y_k(\theta)] = \frac{1}{\pi\sigma^2}\exp\left\{-\frac{u_k(\theta)^2}{\sigma^2}\right\}\exp\left\{-\frac{v_k(\theta)^2}{\sigma^2}\right\}D_k(\theta). \tag{77}$$

Then, by recalling (13) and noticing that $\beta_k = \beta_{k,2q}\beta_{k,2q-1}$, it follows that:

$$p[y_k(\theta)] = \frac{2\beta_{k,2q}}{\sqrt{\pi\sigma^2}}e^{-\frac{u_k(\theta)^2}{\sigma^2}}H_{k,2q} \times\big(u_k(\theta)\big)\frac{2\beta_{k,2q-1}}{\sqrt{\pi\sigma^2}}e^{-\frac{v_k(\theta)^2}{\sigma^2}}H_{k,2q-1}\left(v_k(\theta)\right).$$

By noticing that $p[y_k(\theta)] = p[u_k(\theta), v_k(\theta)]$ and recalling the independence of $u_k(\theta)$ and $v_k(\theta)$, it follows that their pdfs are given by (22) and (23), respectively.

APPENDIX B
PROOF OF (31) AND (32)

• *B.1) Proof of (31):* Before delving into details, we will prove the two following identities:

$$\mathbb{E}\left\{\left[\mathbf{w}(k)^H\dot{\mathbf{a}}(\theta)\right]^2\right\} = 0 \quad \text{and} \quad \mathbb{E}\left\{\left[\dot{\mathbf{a}}(\theta)^H\mathbf{w}(k)\right]^2\right\} = 0. \tag{78}$$

In fact, we have:

$$\begin{aligned}\mathbb{E}\left\{\left[\mathbf{w}(k)^H\dot{\mathbf{a}}(\theta)\right]^2\right\} &= \mathbb{E}\left\{\mathbf{w}(k)^H\dot{\mathbf{a}}(\theta)\mathbf{w}^H(k)\dot{\mathbf{a}}(\theta)\right\},\\ &= \mathbb{E}\left\{\dot{\mathbf{a}}(\theta)^T\mathbf{w}(k)^*\mathbf{w}(k)^H\dot{\mathbf{a}}(\theta)\right\},\\ &= \mathbb{E}\left\{\dot{\mathbf{a}}(\theta)^T\left(\mathbf{w}(k)\mathbf{w}(k)^T\right)^*\dot{\mathbf{a}}(\theta)\right\},\\ &= \dot{\mathbf{a}}(\theta)^T\left(\mathbb{E}\left\{\mathbf{w}(k)\mathbf{w}(k)^T\right\}\right)^*\dot{\mathbf{a}}(\theta).\end{aligned} \tag{79}$$

For *circularly symmetric* noise, however, we have $\mathbb{E}\{\mathbf{w}(k)\mathbf{w}(k)^T\} = \mathbf{0}$ which is used in (79) to obtain the first identity in (78). The second identity in (78) is obtained using equivalent manipulations. Now, applying again the trivial identity $\Re\{z\} = \frac{1}{2}(z + z^*)$ with $z = \dot{\mathbf{a}}(\theta)^H\mathbf{w}(k)$, it follows that :

$$\mathbb{E}\left\{\Re\left\{\dot{\mathbf{a}}(\theta)^H\mathbf{w}(k)\right\}^2\right\} = \frac{1}{4}\mathbb{E}\left\{\left[\mathbf{w}^H(k)\dot{\mathbf{a}}(\theta) + \dot{\mathbf{a}}(\theta)^H\mathbf{w}(k)\right]^2\right\}, \tag{80}$$

and therefore:

$$\begin{aligned}&\mathbb{E}\left\{\Re\left\{\dot{\mathbf{a}}(\theta)^H\mathbf{w}(k)\right\}^2\right\}\\ &= \frac{1}{4}\mathbb{E}\left\{\left[\mathbf{w}^H(k)\dot{\mathbf{a}}(\theta)\right]^2\right\} + \frac{1}{2}\dot{\mathbf{a}}(\theta)^H\mathbb{E}\left\{\mathbf{w}(k)\mathbf{w}^H(k)\right\}\dot{\mathbf{a}}(\theta)\\ &\quad + \frac{1}{4}\mathbb{E}\left\{\left[\dot{\mathbf{a}}(\theta)^H\mathbf{w}(k)\right]^2\right\}.\end{aligned}$$

Then, exploiting the fact that $\mathbb{E}\{\mathbf{w}(k)\mathbf{w}^H(k)\} = \sigma^2\mathbf{I}_{N_a}$ and the two identities in (78), we obtain:

$$\mathbb{E}\left\{\Re\left\{\dot{\mathbf{a}}(\theta)^H\mathbf{w}(k)\right\}^2\right\} = \frac{1}{2}||\dot{\mathbf{a}}(\theta)||^2\sigma^2. \tag{81}$$

• *B.2) Proof of (32):* The expected value of $\Im\{x(k)\}^2$ is obtained by averaging it over all the points in the constellation alphabet:

$$\mathbb{E}\left\{\Im\{x(k)\}^2\right\} = \sum_{c_m\in\mathcal{C}}\Im\{c_m\}^2 Pr[x(k) = c_m]. \tag{82}$$

By using the obvious decomposition $\mathcal{C} = \widetilde{\mathcal{C}} \cup (-\widetilde{\mathcal{C}}) \cup \widetilde{\mathcal{C}}^* \cup (-\widetilde{\mathcal{C}}^*)$ where $\widetilde{\mathcal{C}}$ is the top-right quadrant of the constellation and noticing that $\Im\{\tilde{c}_m\}^2 = \Im\{-\tilde{c}_m\}^2 = \Im\{\tilde{c}_m^*\}^2 = \Im\{-\tilde{c}_m^*\}^2\ \forall\ \tilde{c}_m \in \widetilde{\mathcal{C}}$, (82) can be rewritten as follows:

$$\begin{aligned}&\mathbb{E}\left\{\Im\{x(k)\}^2\right\}\\ &= \sum_{\tilde{c}_m\in\widetilde{\mathcal{C}}}\Im\{\tilde{c}_m\}^2\Big(Pr[x(k) = \tilde{c}_m]\\ &\quad + Pr[x(k) = -\tilde{c}_m] + Pr[x(k) = \tilde{c}_m^*] + Pr[x(k) = -\tilde{c}_m^*]\Big).\end{aligned} \tag{83}$$

Moreover, by using the explicit expressions for the *a priori* probabilities given in Eqs. (37)-(40) of [28], along with the identity $\cosh(x) + \cosh(y) = 2\cosh(\frac{x+y}{2})\cosh(\frac{x-y}{2})$, we obtain:

$$\begin{aligned}&Pr[x(k) = \tilde{c}_m] + Pr[x(k) = -\tilde{c}_m]\\ &+ Pr[x(k) = \tilde{c}_m^*] + Pr[x(k) = -\tilde{c}_m^*]\\ &= 2\beta_k\mu_{k,q}(\tilde{c}_m)\left[\cosh\left(\frac{L_{2q}(k) + L_{2q-1}(k)}{2}\right)\right.\\ &\qquad\left. + \cosh\left(\frac{L_{2q}(k) - L_{2q-1}(k)}{2}\right)\right],\\ &= 4\beta_k\mu_{k,q}(\tilde{c}_m)\cosh\left(\frac{L_{2q}(k)}{2}\right)\cosh\left(\frac{L_{2q-1}(k)}{2}\right).\end{aligned} \tag{84}$$

Since each point $\tilde{c}_m \in \widetilde{\mathcal{C}}_p$ can be written as $\tilde{c}_m = [2n-1]d_q + j[2i-1]d_q$ for some $1 \le i, n \le 2^{q-1}$, the single sum in the right-hand side of (83) can be written as a double sum over the counters n and i. By doing so and using the decomposition in [28], eq. (49)], i.e.:

$$\begin{aligned}\mu_{k,q}(\tilde{c}_m) &= \mu_{k,q}\left([2n-1]d_q + j[2i-1]d_q\right)\\ &= \eta_{k,2q}(i)\eta_{k,2q-1}(n),\end{aligned} \tag{85}$$

it follows that (83) is equivalent to [after using the result of (84)]:

$$\begin{aligned}&\mathbb{E}\left\{\Im\{x(k)\}^2\right\}\\ &= 4\beta_k\sum_{i=1}^{2^{q-1}}\sum_{n=1}^{2^{q-1}}\left[(2n-1)^2d_q^2\,\eta_{k,2q}(i)\eta_{k,2q-1}(n)\right.\\ &\quad\left.\times\cosh\left(\frac{L_{2q}(k)}{2}\right)\cosh\left(\frac{L_{2q-1}(k)}{2}\right)\right]\\ &= \left[2\beta_{k,2q-1}\cosh\left(\frac{L_{2q-1}(k)}{2}\right)\sum_{n=1}^{2^{q-1}}(2n-1)^2d_q^2\eta_{k,2q-1}(n)\right]\\ &\quad\times\left[2\beta_{k,2q}\cosh\left(\frac{L_{2q}(k)}{2}\right)\sum_{i=1}^{2^{q-1}}\eta_{k,2q}(i)\right],\end{aligned} \tag{86}$$

in which the decomposition $\beta_k = \beta_{k,2q}\beta_{k,2q-1}$ was used in the last equality. Moreover, it was recently shown in [26, LEMMA 3] that:

$$2\beta_{k,s}\cosh\left(\frac{L_s(k)}{2}\right)\sum_{n=1}^{2^{q-1}}\eta_{k,s}(n) = 1. \quad (87)$$

By recalling the expression of $\omega_{k,2q-1}$ in (26) and using (87) back into (86), we obtain the result given in (32).

APPENDIX C
EXPRESSION OF $\ddot{u}_k(\theta)$ AND PROOF OF THE INDEPENDENCE OF $\ddot{u}_k(\theta)$ AND $z(k)$

Starting from the expression of $u_k(\theta)$ in (17), we can write:

$$\ddot{u}_k(\theta) = \frac{\partial^2 u_k(\theta)}{\partial\theta^2} = \frac{1}{\sqrt{N_a}}\Re\left\{\ddot{\mathbf{a}}(\theta)^H\mathbf{y}(k)\right\}.$$

Now, replacing $\mathbf{y}(k)$ by its expression, i.e., $S\,\mathbf{a}(\theta)\,x(k)+\mathbf{w}(k)$, it follows that:

$$\ddot{u}_k(\theta) = \frac{S}{\sqrt{N_a}}\Re\left\{\ddot{\mathbf{a}}(\theta)^H\mathbf{a}(\theta)x(k)\right\} + \frac{1}{\sqrt{N_a}}\Re\left\{\ddot{\mathbf{a}}(\theta)^H\mathbf{w}(k)\right\}. \quad (88)$$

Then, by using the identity $\Re\{x\} = \frac{1}{2}(x + x^*)$, we obtain:

$$\ddot{u}_k(\theta) = \frac{1}{2\sqrt{N_a}}\left[S\ddot{\mathbf{a}}(\theta)^H\mathbf{a}(\theta)x(k) + S\mathbf{a}(\theta)^H\ddot{\mathbf{a}}(\theta)x(k)^*\right] + \frac{1}{\sqrt{N_a}}\Re\left\{\ddot{\mathbf{a}}(\theta)^H\mathbf{w}(k)\right\}. \quad (89)$$

Moreover, by recalling the expression of $y_k(\theta)$, it can be shown that $x(k)$ can be expressed as a function of $u_k(\theta)$ and $v_k(\theta)$ as follows:

$$x(k) = \frac{1}{S\sqrt{N_a}}\left[u_k(\theta) + jv_k(\theta)\right] - \frac{1}{SN_a}\mathbf{a}(\theta)^H\mathbf{w}(k). \quad (90)$$

Therefore, plugging (90) back into (89) and using the identity:

$$\ddot{\mathbf{a}}(\theta)^H\mathbf{a}(\theta) + \mathbf{a}(\theta)^H\ddot{\mathbf{a}}(\theta) + 2||\dot{\mathbf{a}}(\theta)||^2 = 0, \quad (91)$$

derived from $\left\|\mathbf{a}(\theta)\right\|^2 = N_a$, it follows that:

$$\ddot{u}_k(\theta) = -\frac{\|\dot{\mathbf{a}}(\theta)||^2}{N_a}u_k(\theta) - \frac{1}{N_a}\Im\left\{\ddot{\mathbf{a}}(\theta)^H\mathbf{a}(\theta)\right\}v_k(\theta) + z_k(\theta),$$

with

$$z_k(\theta) = \frac{1}{\sqrt{N_a}}\Re\left\{\ddot{\mathbf{a}}(\theta)^H\mathbf{w}(k)\right\} - \frac{1}{N_a\sqrt{N_a}}\Re\left\{\ddot{\mathbf{a}}(\theta)^H\mathbf{a}(\theta)\mathbf{a}(\theta)^H\mathbf{w}(k)\right\}.$$

By denoting $\tilde{z}_k(\theta) = \frac{1}{\sqrt{N_a}}\Re\left\{\mathbf{a}(\theta)^H\mathbf{w}(k)\right\}$, it follows that:

$$z_k(\theta) = \ddot{\tilde{z}}_k(\theta) - \frac{1}{N_a\sqrt{N_a}}\Re\left\{\ddot{\mathbf{a}}(\theta)^H\mathbf{a}(\theta)\mathbf{a}(\theta)^H\mathbf{w}(k)\right\}, \quad (92)$$

and

$$u_k(\theta) = S\sqrt{N_a}\Re\{x(k)\} + \tilde{z}_k(\theta). \quad (93)$$

To show the independence of $u_k(\theta)$ and $z_k(\theta)$, notice first that since $z_k(\theta)$ contains only the noise contribution, $\Re\{x(k)\}$ and $z_k(\theta)$ are indeed independent. Moreover, it can be shown that:

$$\begin{aligned}\mathbb{E}\left\{\tilde{z}_k(\theta)z_k(\theta)\right\} &= \mathbb{E}\left\{\tilde{z}_k(\theta)\ddot{\tilde{z}}_k(\theta)\right\} - \frac{1}{N_a\sqrt{N_a}} \\ &\quad\times\mathbb{E}\left\{\tilde{z}_k(\theta)\,\Re\left\{\ddot{\mathbf{a}}(\theta)^H\mathbf{a}(\theta)\mathbf{a}(\theta)^H\mathbf{w}(k)\right\}\right\}, \\ &= \frac{-1}{\sqrt{N_a}}\|\dot{\mathbf{a}}(\theta)\|^2\,\sigma^2 - \frac{1}{\sqrt{N_a}}\Big(-||\dot{\mathbf{a}}(\theta)||^2\sigma^2\Big), \\ &= 0.\end{aligned}$$

Since one can also verify that $\tilde{z}_k(\theta)$ and $z_k(\theta)$ are two Gaussian distributed RVs, it follows that they are independent. Therefore $u_k(\theta)$ and $z_k(\theta)$ are independent. By following the same reasoning, the expression of $\ddot{v}_k(\theta)$, involved in the expression of $\gamma_{k,2q-1}(\theta)$ is obtained as follows:

$$\ddot{v}_k(\theta) = -\frac{||\dot{\mathbf{a}}(\theta)||^2}{N_a}v_k(\theta) + \frac{1}{N_a}\Im\left\{\ddot{\mathbf{a}}(\theta)^H\mathbf{a}(\theta)\right\}u_k(\theta) + z'_k(\theta),$$

where $z_k(\theta)$ is given by:

$$z'_k(\theta) = \frac{1}{\sqrt{N_a}}\Im\left\{\ddot{\mathbf{a}}(\theta)^H\mathbf{w}(k)\right\} - \frac{1}{N_a\sqrt{N_a}}\Im\left\{\ddot{\mathbf{a}}(\theta)^H\mathbf{a}(\theta)\mathbf{a}(\theta)^H\mathbf{w}(k)\right\}.$$

REFERENCES

[1] F. Bellili, C. Elguet, S. B. Amor, S. Affes, and A. Stéphenne, "Closed-form Cramer–Rao lower bounds for DOA estimation from turbo-coded square-QAM-modulated transmissions," in *Proc. IEEE ICASSP*, South Brisbane, QLD, Australia, Apr. 2015, pp. 3492–3496.

[2] H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. New York, NY, USA: Wiley, 2002.

[3] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, Jul. 1996.

[4] A. Hu, T. Lv, H. Gao, Z. Zhang, and S. Yang, "An ESPRIT-based approach for 2-D localization of incoherently distributed sources in massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 996–1011, Oct. 2014.

[5] T. Lv, F. Tan, H. Gao, and S. Yang, "A beamspace approach for 2-D localization of incoherently distributed sources in massive MIMO systems," *Signal Process.*, vol. 121, pp. 30–45, Apr. 2016.

[6] P. Stoica and T. Soderstrom, "Statistical analysis of MUSIC and subspace rotation estimates of sinusoidal frequencies," *IEEE Trans. Signal Process.*, vol. 39, no. 8, pp. 1836–1847, Aug. 1991.

[7] P. Stoica and K. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 38, no. 7, pp. 1132–1143, Jul. 1990.

[8] *Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems-Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands*, IEEE Standard 802.16e-2005, Feb. 2006.

[9] "Technical specification group radio access network; Evolved universal terrestrial radio access (E-UTRA); Physical channels and modulation," Cedex, France, 3GPP TS 36.211, 2010.

[10] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1261–1271, Oct. 1996.

[11] M. Tüchler, R. Koetter, and A. C. Singer, "Turbo equalization: Principles and new results," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 754–767, May 2002.

[12] C. Herzet *et al.*, "Code-aided turbo synchronization," *Proc. IEEE*, vol. 95, no. 6, pp. 1255–1271, Jun. 2007.

[13] N. Noels, H. Steendam, and M. Moeneclaey, "The Cramer–Rao bound for phase estimation from coded linearly modulated signals," *IEEE Commun. Lett.*, vol. 7, no. 5, pp. 207–209, May 2003.

[14] X. Wu and H. Xiang, "Iterative carrier phase recovery methods in turbo receivers," *IEEE Commun. Lett.*, vol. 9, no. 8, pp. 735–737, Aug. 2005.

[15] N. Noels *et al.*, "Turbo synchronization: An EM algorithm interpretation," in *Proc. IEEE Int. Conf. Commun.*, Anchorage, AK, USA, Jun. 2003, pp. 2933–2937.

[16] N. Noels *et al.*, "A theoretical framework for soft-information-based synchronization in iterative (turbo) receivers," *EURASIP J. Wireless Commun. Netw.*, vol. 2005, no. 2, pp. 117–129, Apr. 2005.

[17] V. Lottici and M. Luise, "Embedding carrier phase recovery into iterative decoding of turbo-coded linear modulations," *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 661–669, Apr. 2004.

[18] C. Herzet, X. Wautelet, V. Ramon, and L. Vandendorpe, "Iterative synchronization: EM algorithm versus Newton–Raphson method," in *Proc. IEEE ICASSP*, Toulouse, France, May 2006, pp. 393–396.

[19] 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.

[20] C. Herzet, V. Ramon, and L. Vandendorpe, "Turbo synchronization: A combined sum-product and expectation-maximization algorithm approach," in *Proc. IEEE Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2005, pp. 191–195.

[21] C. Herzet, V. Ramon, L. Vandendorpe, and M. Moeneclaey, "EM algorithm-based timing synchronization in turbo receivers," in *Proc. IEEE ICASSP*, Apr. 2003, pp. 612–615.

[22] V. Ramon, C. Herzet, L. Vandendorpe, and M. Moeneclaey, "EM algorithm-based multiuser synchronization in turbo receivers," in *Proc. IEEE ICASSP*, May 2004, pp. 849–852.

[23] H. Wymeersch and M. Moeneclaey, "Iterative code-aided ML phase estimation and phase ambiguity resolution," *EURASIP J. Appl. Signal Process.*, vol. 2005, no. 6, pp. 981–988, 2005.

[24] L. Zhang and A. G. Burr, "Iterative carrier phase recovery suited to turbo-coded systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 2267–2276, Nov. 2004.

[25] L. Zhang and A. Burr, "A novel carrier phase recovery method for turbo coded QPSK systems," in *Proc. Eur. Wireless (EW) Conf.*, Florence, Italy, 2002, pp. 1–5.

[26] 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.

[27] M. A. Dangl and J. Lindner, "How to use *a priori* information of data symbols for SNR estimation," *IEEE Signal Process. Lett.*, vol. 13, no. 11, pp. 661–664, Nov. 2006.

[28] 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.

[29] M. C. Valenti and B. D. Woerner, "Iterative channel estimation and decoding of pilot symbol assisted turbo codes over flat-fading channels," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 9, pp. 1697–1705, Sep. 2001.

[30] C. Komninakis and R. D. Wesel, "Joint iterative channel estimation and decoding in flat correlated Rayleigh fading," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 9, pp. 1706–1717, Sep. 2001.

[31] R. Otnes and M. Tuchler, "Iterative channel estimation for turbo equalization of time-varying frequency-selective channels," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1918–1923, Nov. 2004.

[32] J. Ylioinas and M. Juntti, "Iterative joint detection, decoding, and channel estimation in turbo-coded MIMO-OFDM," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1784–1796, May 2009.

[33] J.-P. Delmas and H. Abeida, "Cramer–Rao bounds of DOA estimates for BPSK and QPSK Modulated signals," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 117–126, Jan. 2006.

[34] F. Bellili, S. B. Hassen, S. Affes, and A. Stéphenne, "Cramér–Rao lower bounds of DOA estimates from square QAM-modulated signals," *IEEE Trans. Commun.*, vol. 59, no. 6, pp. 1675–1685, Jun. 2011.

[35] A. Leshem and A. J. V. D. Veen, "Direction-of-arrival estimation for constant modulus signals," *IEEE Trans. Signal Process.*, vol. 47, no. 11, pp. 3125–3129, Nov. 1999.

[36] P. Chargé, Y. Wang, and J. Saillard, "An extended cyclic MUSIC algorithm," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1695–1701, Jul. 2003.

[37] P. Chargé and Y. Wang, "A root-MUSIC-like direction finding method for cyclostationary signals," *EURASIP J. Appl. Signal Process.*, vol. 2005, no. 1, pp. 69–73, Jan. 2005.

[38] P. Chargé, Y. Wang, and J. Saillard, "A non-circular sources direction finding method using polynomial rooting," *Signal Process.*, vol. 81, no. 8, pp. 1765–1770, 2001.

[39] R. Krummenauer *et al.*, "Maximum likelihood-based direction-of-arrival estimator for discrete sources," *Circuits Syst. Signal Process.*, vol. 32, no. 5, pp. 2423–2443, Oct. 2013.

[40] R. Attux *et al.*, "A clustering-based method for DOA estimation in wireless communications," in *Proc. 15th Eur. Signal Process. Conf.*, Poznań, Poland, Sep. 2007, pp. 262–266.

[41] M. Pesavento and A. B. Gershman, "Maximum-likelihood direction-of-arrival estimation in the presence of unknown nonuniform noise," *IEEE Trans. Signal Process.*, vol. 49, no. 7, pp. 1310–1324, Jul. 2001.

[42] P. Vallet, P. Loubaton, and X. Mestre, "Improved subspace estimation for multivariate observations of high dimension: The deterministic signals case," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1043–1068, Feb. 2012.

[43] L. Lu and H.-C. Wu, "Novel robust direction-of-arrival-based source localization algorithm for wideband signals," *IEEE Trans. Wireless Commun.*, vol. 11, no. 11, pp. 3850–3859, Nov. 2012.

[44] L. Lu and H.-C. Wu, "Robust expectation–maximization direction-of-arrival estimation algorithm for wideband source signals," *IEEE Trans. Veh. Technol.*, vol. 60, no. 5, pp. 2395–2400, Jun. 2011.

[45] S. A. Vorobyov, A. B. Gershman, and K. M. Wong, "Maximum likelihood direction-of-arrival estimation in unknown noise fields using sparse sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 34–43, Jan. 2005.

[46] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1783–1795, Oct. 1990.

[47] C. E. Chen, F. Lorenzelli, R. E. Hudson, and K. Yao, "Stochastic maximum-likelihood DOA estimation in the presence of unknown nonuniform noise," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3038–3044, Jul. 2008.

[48] A. Renaux, P. Forster, E. Chaumette, and P. Larzabal, "On the high-SNR conditional maximum-likelihood estimator full statistical characterization," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4840–4843, Dec. 2006.

[49] Z.-M. Liu, Z.-T. Huang, and Y.-Y. Zhou, "An efficient maximum likelihood method for direction-of-arrival estimation via sparse Bayesian learning," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 1–11, Oct. 2012.

[50] T. Li and A. Nehorai, "Maximum likelihood direction finding in spatially colored noise fields using sparse sensor arrays," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1048–1062, Mar. 2011.

[51] P. Stoica and P. Babu, "The Gaussian data assumption leads to the largest Cramér–Rao bound [lecture notes]," *IEEE Signal Process. Mag.*, vol. 28, no. 3, pp. 132–133, May 2011.

[52] S. Park, E. Serpedin, and K. Qaraqe, "Gaussian assumption: The least favorable but the most useful [lecture notes]," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 183–186, May 2013.

[53] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, vol. 1. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.

[54] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, no. 5, pp. 5–14, May 1990.

[55] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications," *IEEE Signal Process. Mag.*, vol. 17, no. 3, pp. 29–48, May 2000.

[56] N. Noels, H. Steendam, and M. Moeneclaey, "Carrier and clock recovery in (turbo-) coded systems: Cramér–Rao bound and synchronizer performance," *EURASIP J. App. Signal Process.*, vol. 2005, no. 6, pp. 972–980, May 2005.

[57] H. Wang, S. Kay, and S. Saha, "An importance sampling maximum likelihood direction of arrival estimator," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5082–5092, Oct. 2008.

[58] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.

[59] M. M. Mansour and N. R. Shanbhag, "High-throughput LDPC decoders," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 11, no. 6, pp. 976–996, Dec. 2003.

[60] M. Valenti. (2008). A guided tour of CML, the coded modulation library. Iterative Solutions, Morgantown, WV, USA. [Online]. Available: http://www.iterativesolutions.com/user/image/cmloverview.pdf

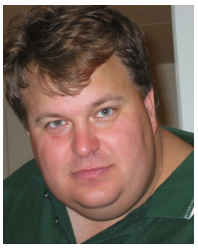
Faouzi Bellili was born in Tunisia in 1983. He received the B.Eng. degree (Hons.) in signals and systems from the Tunisia Polytechnic School in 2007, and the M.Sc. and Ph.D. degrees (Hons.) from the National Institute of Scientific Research (INRS-EMT), University of Québec, Montreal, QC, Canada, in 2009 and 2014, respectively. From 2014 to 2016, he was a Research Associate with INRS-EMT, where he coordinated a major multi-institutional NSERC Collaborative Research and Development Project on 5th-Generation (5G) Wireless Access Virtualization Enabling Schemes (5G-WAVES), a key tech derive of the Horizon 2020 worldwide economic and social development initiative. He is currently a Post-Doctoral Fellow with the University of Toronto, ON, Canada. His research focuses on statistical and array signal processing with an emphasis on parameters estimation for wireless communications and 5G-enabling technologies. He has authored/co-authored over 50 peer-reviewed papers in reputable IEEE journals and conferences.

Dr. Bellili received the Best M.Sc. Thesis Award from INRS-EMT for the year 2009–2010 and the National Grant of Excellence from the Tunisian Government twice for the M.Sc. and Ph.D. programs. In 2011, he also received the Merit Scholarship for Foreign Students from the Ministère de l'Éducation, du Loisir et du Sport of Québec, Canada, for the period 2011–2013. He was also selected by INRS as its candidate for the 2009–2010 competition of the very prestigious Vanier Canada Graduate Scholarships program. He also received the very prestigious Academic Gold Medal of the Governor General of Canada for the year 2009–2010 and the Excellence Grant of the Director General of INRS for the year 2009–2010. In 2015, he was also selected by INRS as its candidate for the 2014–2015 competition of the very prestigious Doctoral Dissertation Award of the Northeastern Association of Graduate schools, the University of Waterloo and the University of Toronto as one of their candidates to the very prestigious Banting Post-Doctoral Fellowship competition for the year 2015/2016 in 2015. He received the INRS Innovation Award for the year 2014/2015. He also received the prestigious PDF Scholarship offered over the same period (but declined) from the Fonds de Recherche du Québec Nature et Technologies. He received the very prestigious NSERC PDF Grant for the period 2017–2018. He serves regularly as a TPC member for major IEEE conferences and acts as a reviewer for many international scientific journals and conferences.

Chaima Elguet was born in Tunisia in 1990. She received the Dipl.Ing. degree in telecommunications from The Higher School of Communication of Tunis, Tunisia, in 2014, and the M.Sc. degree from the University of Québec (INRS-EMT), Montreal, Canada, under the advisement of Professor S. Affes in 2016. Her research works are mainly focused in the areas of wireless communication theory.

Souheib Ben Amor was born in Hammam Sousse, Sousse, Tunisia, in 1989. He received the Dipl.Ing. degree in telecommunications from the National Engineering School of Tunis in 2013 and the M.Sc. degree from the Institut National de la Recherche Scientifique-Énergie, Matériaux, et Télécommunications (INRS- EMT), Université du Québec, Montréal, QC, Canada, in 2016, where he is currently pursuing the Ph.D. degree, with a focus on statistical signal processing and array processing and their applications in wireless communications. He received a scholarship for the M.Sc. degree under an agreement between the Tunisian government and INRS-ÉMT. He also received the INRS scholarship for the Ph.D. degree.

Sofiène Affes (S'94–SM'04) received the Dipl.Ing. degree in telecommunications and the Ph.D. degree (Hons.) in signal processing from the École Nationale Supérieure des Télécommunications, Paris, France, in 1992 and 1995, respectively. He was a Research Associate with INRS, Montreal, QC, Canada, until 1997, an Assistant Professor until 2000, and an Associate Professor until 2009. He is currently a Full Professor and the Director of PERSWADE, a unique U.S. $4 million research training program on wireless in Canada, involving 27 faculty members from 8 universities and 10 industrial partners. From 2003 to 2013, he was a Canada Research Chair in wireless communications. He has been a recipient of a Discovery Accelerator Supplement Award twice from NSERC from 2008 to 2011 and from 2013 to 2016. He is currently an Associate Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS and the *Journal on Wireless Communications and Mobile Computing* (Wiley). He was an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He served as a General Co-Chair of the IEEE VTC2006, Montreal, QC, Canada, and the IEEE ICUWB 2015, Montreal. For his contributions to the success of both events, he received a Recognition Award from the IEEE Vehicular Technology Society in 2008 and a Certificate of Recognition from the IEEE Microwave Theory and Techniques Society in 2015. He is currently serving as the General Chair of 28th IEEE PIMRC to be held in Montreal in 2017.



Alex Stéphenne (M'95–SM'06) was born in Quebec, Canada, in 1969. He received the B.Ing. degree in electrical engineering from McGill University, Montreal, Quebec, in 1992, and the M.Sc. and Ph.D. degrees in telecommunications from INRS-Télécommunications, Université du Québec, Montreal, in 1994 and 2000, respectively. In 1999, he joined SITA Inc., Montreal, where he was involved in the design of remote management strategies for the computer systems of airline companies. In 2000, he became a DSP Design Specialist for Dataradio Inc., (part of CalAmp), Montreal, a company specializing in the design and manufacturing of advanced wireless data products and systems for mission critical applications. In 2001, he joined Ericsson and worked for over two years in Sweden, where he was responsible for the design of baseband algorithms for WCDMA commercial base station receivers. From 2003 to 2008, he was with Ericsson, Montreal, where he was a Researcher with a focus on issues related to the physical layer of wireless communication systems. From 2009 to 2014, he was a Researcher with Huawei, Ottawa, where he was involved in radio resource management for 4G and 5G systems. Since 2014, he has been a Wireless System Designer with Ericsson, Ottawa. He has also been an Adjunct Professor with INRS since 2004.