Cramér-Rao Bounds for spectral parametric estimation with compressive multiband architectures

Marguerite Marnat, Michael Pelissier, Laurent Ros, Olivier J.J. Michel

Cramér-Rao Bounds for spectral parametric estimation with compressive multiband architectures

Marguerite Marnat*, Michaël Pelissier*, Laurent Ros[†] and Olivier Michel [†]

*Univ. Grenoble Alpes, CEA, LETI, F-38000 Grenoble

[†] Univ. Grenoble Alpes, CNRS, Grenoble INP[‡], GIPSA-lab, 38000 Grenoble, France

[‡]Institute of Engineering Univ. Grenoble Alpes

Email: olivier.michel@grenoble-inp.fr

Abstract—This article tackles the topic of performance analysis for Spectrum Sensing based on Compressive Sampling (CS). More precisely, the lower bound on the variance of any unbiased estimator, the Cramér-Rao Bound (CRB), is investigated in the context of spectral parametric estimation. Compressed samples are obtained from a multiband architecture like the Modulated Wideband Converter, the Quadrature-Analog-to-Information Converter or the Periodic Non Uniform Sampler. An expression of the Fisher information matrix, which allows to compute the CRB, is established for a compressive multiband architecture assuming a disjoint spectral subband model. The relationships between Fisher matrices in a generic framework are exposed: first between compressive multiband and subsampling architectures, then between subsampling and Nyquist sampling architectures. Based on these new considerations, the issue of interferer detection, a canonical case and also a huge thorn in the side of wideband radiofrequency receivers is tackled. A benchmark between Nyquist, Subsampling and Compressive Multiband Sampling approaches is provided for frequency and amplitude estimation of dual-tone signals. This analysis illustrates the way in which interferences between parameters occur in estimation with Compressive Sampling. It is then shown how properties of the sensing matrix for popular compressive architectures enable to control the precision of spectral parametric estimation in specific subbands. This control opportunity opens the door to adaptive methods.

Index Terms—Compressive Sensing, Parametric estimation, Cognitive Radio, Cramér-Rao Bound, Modulated Wideband Converter, Interferer detection

I. INTRODUCTION

THE Compressive Sensing (CS) theory introduced in 2006 by Candès *et al.* [1] represents a paradigm shift able to cope with the acquisition rate deadlock presented by the Nyquist rate. By focusing instead on the concept of information, CS is foreseen to disrupt the usual trade-off between bandwidth, noise figure and energy consumption in Analog-to-Digital Converters conception. The key assumption for its applicability is sparsity, the existence of a representation space in which few coefficients are sufficient to describe the signal. This is true for the radiofrequency spectrum which is mostly empty at a given time and space, meaning that the wideband radiofrequency signal is sparse in the frequency domain (with e.g. a Fourier matrix as the sparsity matrix). Instead of acquiring redundancies at Nyquist rate, only the relevant information is captured and the sparsity hypothesis suffices to enable recovery if enough diversity is created

between the measures.

However the complete reconstruction of the signal is a non-linear optimization process, inducing high computational costs and hence not compatible with the context of embedded radio solutions. Yet reconstruction is not always necessary and a promising remedy to this drawback is to perform the extraction of the features of interest directly from the compressed samples. Due to its low computational cost, this approach is relevant for various applications. The authors of [2] detailed in particular the potential for Cognitive Radios (CR), radios that are configured dynamically based on available information and which only require a certain understanding of the spectral environment. Applications described in [2] include wideband Spectrum Sensing (energy level, power spectrum, ...), wireless channel parameter estimation (channel state information, direction of arrival, signal to noise ratio, sparsity order, ...) and radio environment map construction. Proposed wideband compressive radiofrequency (RF) receivers for CR applications (Modulated Wideband Converter [3], Periodic Non Uniform Sampler [4], Nyquist Folding Receiver [5], Non Uniform Wavelet Bandpass Sampling [6]) are essentially based on controlled spectrum folding operations.

A critical issue in compressive Spectrum Sensing, which is pinpointed by [7] and [8], is that wideband compressive architectures must account for possible interferences across the whole input range and hence strong out-of-band blockers limit the dynamic range. Easily detecting these blockers and potentially taming their effect is therefore a topic of high interest on which we will focus the application case.

Compared to parametric estimation based on samples acquired at the Nyquist frequency, parametric estimation based on compressed samples is a domain of growing interest. A convenient performance evaluation tool is the Cramér-Rao lower Bound (CRB) on the variance of any unbiased estimator. The CRB on the variance of the estimated coordinates of a sparse parameter vector was derived in [9], which is of major interest for signal reconstruction in the CS framework. In [10], a parametric estimation based on compressed samples is developed, with a specific focus on frequency estimation. Bounds are derived and asymptotic results are given as a function of the statistical properties of the sensing matrix. However in order to size an architecture according to specifications, a more thorough framework is needed: with asymptotic constraints on each parameter to estimate instead of global guarantees, and with

bounds straightforwardly computable from the coefficients of the sensing matrix instead of statistical guarantees.

In this paper, a framework for parametric estimation based on the samples of a multiband compressive RF receiver is presented for interferer detection applications. Assuming a disjoint subband spectrum model, the Cramér-Rao Bound for each parameter is derived based on the deterministic sensing matrix. This assumption enables the estimation of performance in this context to be more accurate than aforementioned results. The closed form expression established in a generic framework is then illustrated in a concrete case to highlight the role of interferences in parametric estimation from compressed samples.

This article is organized as follows. In Section II, notations and framework are introduced. In Section III, a State-of-the-Art of compressive parameter estimation and the corresponding Cramér-Rao Bound is sketched. The next two Sections contain our contributions. In Section IV, a new theoretical result on the Fisher information matrix with samples from a compressed multiband architecture is established assuming a spectral disjoint subband model. In Section V, the interferer detection issue is tackled through two detailed and benchmarked application cases: for a real dual-tone signal, amplitude estimation then frequency estimation are carried out. In Section VI, conclusions and perspectives are drawn.

II. FRAMEWORK

A notation table is provided in Appendix E. Matrices and vectors will be written with uppercase and lowercase boldface letters respectively. $\lfloor . \rfloor$ will denote the floor operator, $\mod (.)$ the modulo operator, $.^*$ the conjugate, $.^H$ the conjugate transpose and $[\![\ldots ; \ldots]\!]$ an interval of integers. The tilde $\tilde{.}$ will refer to the continuous Fourier Transform (of an analog or discrete time signal). Indices of a vector are noted with $[.]$ in both time and frequency domains

A. Compressive Sensing

Let $\boldsymbol{x} \in \mathbb{C}^{\nu}$ be a vector of length ν containing the time series or sampled representation of a continuous complex signal $x(t)$. In Compressive Sensing $\boldsymbol{x}$ is assumed to be sparse with respect to $\boldsymbol{\Psi} \in \mathbb{C}^{\nu \times \xi}$ in the sense that it exists $\boldsymbol{s}$ such that $\boldsymbol{x} = \boldsymbol{\Psi s}$ with $||\boldsymbol{s}||_0 \ll \nu$, i.e. $\boldsymbol{s}$ has few non zero components. The columns of $\boldsymbol{\Psi}$ represent dictionary elements and in general $\xi \geq \nu$. Let $\boldsymbol{\Phi}$ denotes the acquisition matrix acting on the noisy input signal $\boldsymbol{x}^+ = \boldsymbol{x} + \boldsymbol{w}$:

$$\boldsymbol{y}^+ = \boldsymbol{\Phi}(\boldsymbol{x} + \boldsymbol{w}) = \boldsymbol{y} + \boldsymbol{w}_{CS} \qquad (1)$$

$$\boldsymbol{y} = \boldsymbol{\Phi \Psi s} \qquad (2)$$

where $\boldsymbol{w}$ and $\boldsymbol{w}_{CS} = \boldsymbol{\Phi w}$ are respectively the noise in input and after compression and $\boldsymbol{y}$ and $\boldsymbol{y}^+$ the noiseless and noisy output signal. It is assumed that $\boldsymbol{\Phi} \in \mathbb{C}^{\mu \times \nu}$ and $\mu < \nu$. In this paper focused on Cognitive Radio applications, $\boldsymbol{x}$ will be assumed to be sparse in the frequency domain. Thus $\boldsymbol{x}$ is the signal at the receiver input sampled at Nyquist rate $f_{Nyq} = \frac{1}{T_{Nyq}}$ during the acquisition time $T_{acq} = \nu.T_{Nyq}$, $\xi = \nu$, $\boldsymbol{\Psi}$ is the Discrete Fourier Transform (DFT) matrix $\boldsymbol{F}$ with generic term $[F]_{k,i} = e^{-j2\pi ki}$ and $\boldsymbol{s} = \tilde{\boldsymbol{x}} = DFT(\boldsymbol{x})$.
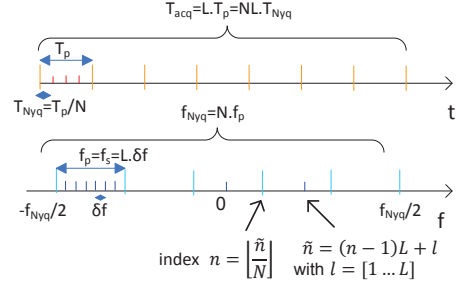

Fig. 1: Time and frequency correspondence.

We will consider that the Nyquist band defined by the receiver $[-\frac{f_{Nyq}}{2}; \frac{f_{Nyq}}{2}]$ is split into N subbands so that $f_{Nyq} = N.f_p$, where $f_p = \frac{1}{T_p}$ is the subband spacing. The acquisition length is then set to be $\nu = NL$ where L is the number of bins per subband. The frequency resolution is thus given by $\delta f = \frac{f_{Nyq}}{NL}$. The output dimension is $\mu = ML$ where M is typically the number of channels. In this paper the vector $\tilde{\boldsymbol{y}}$ concatenates the spectrum components of the M acquisition channels, each channel having L components. It will be useful to introduce the global frequency index $\tilde{n} = (n-1)L + l$, to index the bin l in the frequency subband n, where $n \in [\![1; N]\!]$ and $l \in [\![1; L]\!]$. An illustration of time frequency correspondence is given in Fig. 1. In this study, the number K of active frequency subbands will be assumed to satisfy $K \ll N$ and will abusively be referred to as the sparsity degree. The frequency support of $\boldsymbol{x}$ is given by:

$$\Lambda = \{n \in [\![1; N]\!] | \ \exists (n,l) \text{ satisfying } |\boldsymbol{s}[\tilde{n}]| > 0\} \qquad (3)$$

$\Lambda(f)$ denotes the index n corresponding to the subband containing frequency f. Although there is no straightforward way to evaluate the quality of a sensing matrix, norm and orthogonality preservation are two helpful metrics.

- A matrix $\boldsymbol{\Phi}$ is said to satisfy the Restricted Isometry Property (RIP) [1] with parameters (K, δ_K) if:

$$(1 - \delta_K) \|\boldsymbol{\Psi s}\|_2^2 \leq \|\boldsymbol{\Phi \Psi s}\|_2^2 \leq (1 + \delta_K) \|\boldsymbol{\Psi s}\|_2^2 \quad (4)$$

  for all K-sparse vectors $\boldsymbol{x} = \boldsymbol{\Psi s} \in \mathbb{C}^{\nu}$. The RIP measures the norm deformation by the projection.

- The coherence [1] of the matrix $\boldsymbol{B} = \boldsymbol{\Phi \Psi}$ is the largest absolute Hermitian inner product between any two different normalized columns:

$$coherence(\boldsymbol{B}) = \max_{i \neq j} \left(\frac{|<\boldsymbol{B}_{.,i}, \boldsymbol{B}_{.,j}>|}{\|\boldsymbol{B}_{.,i}\| \|\boldsymbol{B}_{.,j}\|} \right) \quad (5)$$

  where $\boldsymbol{B}_{.,i}$ is the ith column of $\boldsymbol{B}$. Coherence measures a deviation from the orthogonality condition between the columns of the matrix. This quasi orthogonality ensures that the input information is spread among all the measurements.

- The Gramian of a matrix $\boldsymbol{B}$ is given by $\boldsymbol{B}^H \boldsymbol{B}$. Note that for a normalized sensing matrix (with unitary column energy), the maximum magnitude of the Gramian outside the diagonal corresponds to the coherence. The K-RIP property is given by extremal eigenvalues over all support Λ of cardinal K.

If the input noise is assumed to be a white (complex circular) Gaussian noise, $\boldsymbol{w} \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I}_{NL})$, the noise after compression is then $\boldsymbol{\Phi} \boldsymbol{w} \sim \mathcal{N}(0, \sigma^2 \boldsymbol{\Phi} \boldsymbol{\Phi}^H)$. If $\boldsymbol{\Psi}$ is orthogonal and $\boldsymbol{\Phi}$ satisfies to an acceptable degree the properties of coherence and RIP, it is shown in [11] that assuming $\boldsymbol{\Phi} \boldsymbol{w} \sim \mathcal{N}\left(0, \frac{N}{M} \sigma^2 \boldsymbol{I}_{ML}\right)$ is an accurate approximation. Hence in the rest of the paper the model of $\boldsymbol{w}_{CS}$ in eq. (1) is assumed to be white Gaussian:

$$\boldsymbol{w}_{CS} \sim \mathcal{N}\left(0, \frac{N}{M} \sigma^2 \boldsymbol{I}_{ML}\right) \qquad (6)$$

B. Compressive multiband architectures

Multiband architectures will refer to devices that process identically all frequency bins within the same subband. Then instead of considering the architecture-agnostic $\boldsymbol{B}$, the acquisition operations between the input and output spectra

$$\tilde{\boldsymbol{y}} = \check{\boldsymbol{B}} \tilde{\boldsymbol{x}} \qquad (7)$$

can be described, as illustrated in Fig. 2a, by a $M.L \times N.L$ block matrix $\check{\boldsymbol{B}} = \boldsymbol{B} \otimes \boldsymbol{I}_L$ where $\otimes$ stands for the Kronecker product, $\boldsymbol{B} \in \mathbb{C}^{M \times N}$ and $\boldsymbol{I}_L$ is the identity matrix of size $L \times L$. This formalism will be helpful for generalized demonstrations in the next sections. In order to get a multiband sensing matrix that does not depend on the acquisition length NL, it is usually preferred as in [3] to consider the sensing matrix $\boldsymbol{B} \in \mathbb{C}^{M times N}$, and $\tilde{\boldsymbol{x}}_\sharp$ and $\tilde{\boldsymbol{y}}_\sharp$ so that input $\tilde{\boldsymbol{x}}$ is reordered in a matrix $\tilde{\boldsymbol{x}}_\sharp$ of N rows of L bins, and output $\tilde{\boldsymbol{y}}$ is reordered in a matrix $\tilde{\boldsymbol{y}}_\sharp$ of M rows of L bins. $\boldsymbol{B}$ is illustrated in Fig. 2b. As the considered frequency band is zero-centered, it finally yields, for odd N and $(m, n, l) \in [\![1; M]\!] \times [\![1; N]\!] \times [\![1; L]\!]$:

$$x_\sharp[n, l] = \tilde{x}\left(\left(n - \lfloor \tfrac{N}{2} \rfloor - \tfrac{3}{2}\right) f_p + (l-1)\delta f\right) \qquad (8)$$

$$y_\sharp[m, l] = \tilde{y}_m\left((l-1)\delta f - \frac{f_p}{2}\right) \qquad (9)$$

Then eq. (7) transforms into the simpler formalism:

$$\tilde{\boldsymbol{y}}_\sharp = \boldsymbol{B} \tilde{\boldsymbol{x}}_\sharp \qquad (10)$$

Multiband architectures are particularly well suited for multiband signals but also for multitones. A typical example of those architectures is the Modulated Wideband Converter (MWC) introduced by Mishali and Eldar [3], pictured in Fig. 3. In a MWC the input signal is mixed in each of the M parallel branches with T_p-periodic functions. For each branch $m \in [\![1; M]\!]$, the mixing function $p_m(t)$ consists, within each period, of a code $\boldsymbol{\alpha}_m$ of N elements shaped by e.g. rectangular chip pulses. As seen in Fig. 3, in the frequency domain this amounts to a convolution with a f_p-spaced Dirac comb distribution so that each band is weighted by the corresponding Fourier coefficient of the code and the whole spectrum is aliased at baseband. The next step consists in low-pass filtering with cut-off frequency $f_c = 1/2T_s$ and uniform sampling at f_s (by default $f_s = 1/T_s = f_p$). Appendix A demonstrates anew that for a Modulated Wideband Converter $b_{m,n}$ corresponds to the Fourier coefficients of the time domain mixing codes $p_m(t)$ in the mth branch,
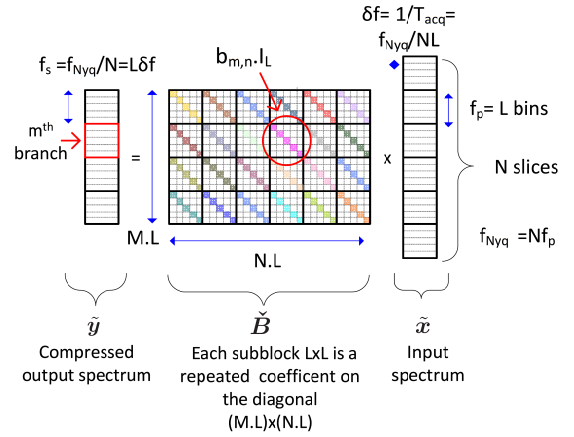
(a) Multiband architecture, formalism $\check{\boldsymbol{B}}$.
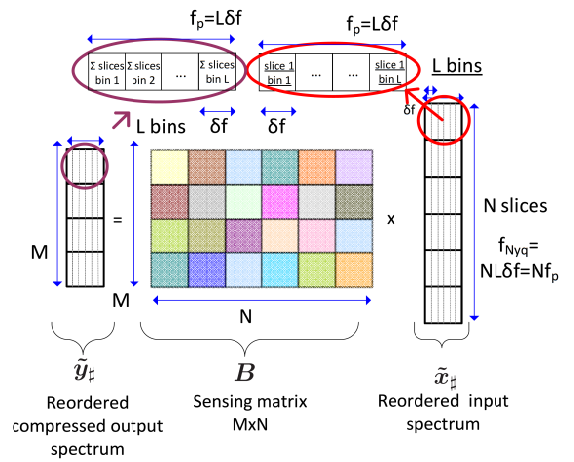
(b) Multiband architecture, formalism $\boldsymbol{B}$.

Fig. 2: Sensing matrix of a multiband architecture for two different formalisms.
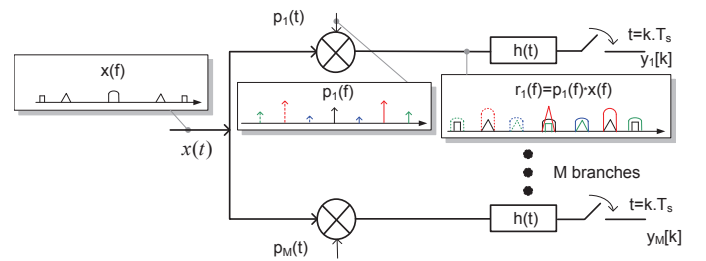
Fig. 3: Block diagram of a Modulated Wideband Converter Architecture.

indexed by $n - \lfloor \frac{N}{2} \rfloor - 1$. Bandpass Sampling (BP) [12] which is a controlled undersampling causing folding of multiband components into the lowpass region also known as harmonic sampling, Multirate Sampling (MRS) [13] and Periodic Non Uniform Sampling (PNUS) [4] can be seen as particular cases of multiband architectures and fit under the same multiband

formalism, with different matrices B:

$$\text{For Bandpass: } b_{1,n} = 1 \ (M{=}1) \tag{11}$$

$$\text{For PNUS: } b_{m,n} = e^{-2j\pi np[m]\delta f} \tag{12}$$

$$\text{For MRS: } b_{m,n} = 1, \text{variable } L \text{ in each branch} \tag{13}$$

where $p[m] \in [\![1; N]\!]$ is a pattern of M among N samples. The compression ratio is given by M/N, so that at fixed number of samples the equivalent observable bandwidth is N/M times wider.

The Quadrature-Analog-to-Information-Converter (QAIC) proposed in [14] is a variation of the MWC that performs frequency down-conversion of the real-valued signal into a baseband complex signal before analog mixing with the pseudorandom codes. The QAIC showed appealing performances for interferer detection [15]. To generalize the results to downconverted signals, both complex and real-valued cases for signal x will be considered. Our working scenario is focused on interferer detection near a useful signal and hence amplitude and frequency estimation of a dual-tone signal. To this end we will first introduce in the next Subsection the Cramér-Rao lower Bound on unbiased estimator variance.

C. Cramér-Rao Bound

Let us note with $\hat{\cdot}$ an estimator. The amount of information on parameters which is carried by a signal observation can be measured by the Fisher information matrix, noted J. More precisely, let us note with $\Gamma \in \mathbb{R}^{Q \times Q}$ the covariance matrix of an unbiased estimator of a $Q \times 1$ parameter vector θ. Then under mild conditions [16], Γ is lower bounded, in the sense of positive definiteness, by the Cramér-Rao lower Bound (CRB) [16]:

$$\Gamma \geq CRB(\theta) = J(\theta)^{-1} \tag{14}$$

$$\Gamma - J(\theta)^{-1} \geq 0 \tag{15}$$

Therefore the unbiased estimation of θ_q, the qth parameter of the vector θ ($q \in [\![1; Q]\!]$), is lower bounded by [16]:

$$var(\hat{\theta}_q) \geq CRB(\hat{\theta}_q) = (J(\theta)^{-1})[q, q] \geq (J(\theta)[q, q])^{-1} \tag{16}$$

where the two right hand terms are equal if and only if J is diagonal. If $J[q, r] \neq 0$ for $q \neq r$, the additional parameter interferes, degrading performances.

Note that for a complex-valued vector $x^+[\tilde{n}; \theta] = x[\tilde{n}; \theta] + w[\tilde{n}]$ where $w \sim \mathcal{N}(0, \sigma^2 I_{NL})$ is a circular white complex Gaussian noise [[16], Chapter 15], for $(q, r) \in [\![1; Q]\!]^2$:

$$J[q, r] = \frac{2}{\sigma^2} \Re \left[\sum_{\tilde{n}=1}^{NL} \frac{\partial x^*[\tilde{n}; \theta]}{\partial \theta_q} \frac{\partial x[\tilde{n}; \theta]}{\partial \theta_r} \right] \tag{17}$$

where $\Re[.]$ denotes the real part. In the next Section the current State-of-the-art of Cramér-Rao Bounds applied to parametric estimation based on compressed samples is presented, and the importance of the choice of the sensing matrix is highlighted.

III. PARAMETRIC ESTIMATION WITH COMPRESSED SAMPLES: STATE OF THE ART

As far as spectral information is concerned, the estimation of the parameters governing the underlying signal model is mostly based on Nyquist rate sampling since folding must be avoided to preserve the information. Reconstruction relies on costly and difficult non convex optimization methods. It is proposed here to circumvent this problem and extract parameters directly from the compressed samples. This approach differs strongly from traditional parameter estimation in the sense that the structure of the information has been scrambled along in the acquisition process. When reconstruction is performed after the compressive sensing acquisition, an additional assumption on the sparsity of the input enables to preserve the information. However extracting spectral information directly from the compressed samples is another challenge.

A. State-of-the-art

Signal reconstruction can be viewed as a specific case of parameter estimation with $\theta = x$. The authors of [9] showed that, for known signal support Λ or unknown but with cardinality exactly equal to the maximum sparsity level, the Cramér-Rao Bound of $x[\tilde{n}]$ for $\tilde{n} \in [\![1; NL]\!]$ is given by:

$$CRB(x) = E\left(\|x - \hat{x}\|_2^2\right) = \sigma^2 Tr\left(\Phi_\Lambda^H \Phi_\Lambda\right) \tag{18}$$

where Φ_Λ is the restriction of the sensing matrix to the support Λ and Tr(.) is the trace operator.

For parametric estimation, an interesting framing of the Fisher matrix has been established in [10], [17] by approaching the limit on the distance preservation property ($2K$-RIP), thus extending norm preservation to the partial derivative vector. And since the Fisher matrix is the Gramian of the partial derivative vector:

$$\frac{M}{N}(1 - \delta_{2K})^2 J_{Nyq} \leq J_{CS} \leq \frac{M}{N}(1 + \delta_{2K})^2 J_{Nyq} \tag{19}$$

where J_{Nyq} and J_{CS} are the Fisher information matrices given the Nyquist or the compressed samples respectively. $\frac{M}{N}$ conforms the noise folding aspect raised in (6). $(1 \pm \delta_{2K})^2$ corresponds to the unavoidable non-isometry of the projection (RIP). An example case is given for frequency estimation of a dual-tone signal, where the bound becomes $O(\frac{\sigma^2}{N^2 M})$ for compressive sampling instead of $O(\frac{\sigma^2}{N^3})$. Some more specific analyses were derived for parametric model estimation with random distributions, for example the mean of a complex distribution for direction-of-arrival estimation in [18].

B. Positioning of our approach

Previous works ([10], [18]) consider a random matrix with guarantees relating to the matrix as a whole from a statistical point of view. Yet in practice, sensing matrices, whether randomly generated from a distribution or structured, are deterministic. The purpose of this article is to explore more precisely how each one of the parameter estimates is specifically affected by the deterministic coefficients of the compressive sensing matrix, taking into account the knowledge of the acquisition process in a concrete application scenario.
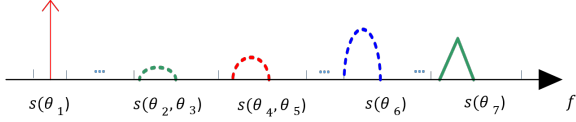
Fig. 4: Disjoint subband model.

In the next Section, a new theorem on Cramér-Rao Bounds with compressive multiband architectures is established. Relationships between Fisher information matrices for Compressive Multiband Sensing, Subsampling and Nyquist Sampling will be exposed.

IV. CRAMÉR-RAO BOUND FOR MULTIBAND ARCHITECTURES

Consider a $Q \times 1$ parameter vector $\boldsymbol{\theta}$ to be estimated and a K-sparse signal $\boldsymbol{x}$, containing exactly K active subbands. If there is a unique unknown parameter in each active subband, then $K = Q$. In general $K \neq Q$.

Assumption 1 (Disjoint subband model) *Assume the existence of an underlying "disjoint" signal model, meaning that each parameter θ_q ($q \in [\![1; Q]\!]$) does not influence more than one subband signal and no subband signal spread over multiple subbands. This translates into the existence of a function g that maps each parameter θ_q with its unique corresponding support element $g(\theta_q) = \lambda_q \in \boldsymbol{\Lambda}$ such that: $\forall q \in [\![1; Q]\!]$ if $\lambda_q \neq \Lambda(f)$ then*

$$\frac{\partial \tilde{x}(f)}{\partial \theta_q} = 0 \qquad (20)$$

An example is given in Fig. 4. Tones satisfy this model and so do modulated signals whose channel bandwidths belong to a unique frequency subband of the architecture. This would not be the case for more wideband signals. For sake of conciseness, $\tilde{x}_{\lambda_q}$ denotes the component of the signal influenced by parameter q and its central frequencies will be noted f_k for $k \in [\![1; K]\!]$.

Theorem 1 (Multiband architectures) *Assuming the above model, and a real-valued signal $\boldsymbol{x}$, the Fisher information matrix is given by:*

$$\boldsymbol{J}_{MB} = \Re\left[\boldsymbol{B}_{\boldsymbol{\Lambda}}^H \boldsymbol{B}_{\boldsymbol{\Lambda}}\right] \circ \boldsymbol{J}_{BP} \qquad (21)$$

where $\boldsymbol{J}_{MB}$ and $\boldsymbol{J}_{BP}$ are the Fisher information matrices for Compressive Sensing acquisition with a multiband architecture, and Bandpass Sampling respectively. $\boldsymbol{B}_{\boldsymbol{\Lambda}}$ is the column restriction of the matrix $\boldsymbol{B}$ to the support $\boldsymbol{\Lambda}$ with possible multiplicity of a subband index for different parameters and is normalized w.r.t. column energy, $\Re[.]$ the real part and $\circ$ is the Hadamard (elementwise) product.
If the signal $\boldsymbol{x}$ is complex-valued eq. (21) transforms into:

$$\boldsymbol{J}_{MB} = \Re\left[\boldsymbol{B}_{\boldsymbol{\Lambda}}^H \boldsymbol{B}_{\boldsymbol{\Lambda}}\right] \circ \boldsymbol{J}_{BP} + \Im\left[\boldsymbol{B}_{\boldsymbol{\Lambda}}^H \boldsymbol{B}_{\boldsymbol{\Lambda}}\right] \circ \check{\boldsymbol{J}}_{BP} \qquad (22)$$

where $\Im[.]$ denotes the imaginary part. $\check{J}_{BP}[q, r]$ corresponds to the $[q, r]^{th}$ coefficient of the Fisher matrix for Bandpass Sampling in the case where the initial dephasing between

components x_{λ_q} and x_{λ_r} is increased by $\frac{\pi}{2}$ compared to the actual input signal.

Refer to Appendix B for a proof.

Assumption 2 (Bandlimited signal) *Assume that Asm. A1 is satisfied and also that the spectral representation of the samples of each subsignal covers at most L bins:*

$$\forall q \in [\![1; Q]\!] \ \tilde{x}_q[\tilde{n}] = 0 \ for \ \tilde{n} \notin [\![\tilde{n}_0(q) - \frac{L}{2}; \tilde{n}_0(q) + \frac{L}{2}]\!] \ (23)$$

where $\tilde{n}_0(q)$ corresponds to the central bin of $\tilde{x}_q$. Eq. (23) expresses the fact that there is no leakage between subbands.

Due to finite duration of the acquisition and therefore spectral leakage, this hypothesis is an approximation. Note that Asm. A2 deals, through the decomposition of the signal on the frequency grid, with bandlimited discretized representations whereas Asm. A1, in contrast, was dealing with bandlimited analog representations and hence did not need to neglect spectral leakage on remote frequencies.

Theorem 2 (Multiband architectures) *Under the above assumptions A1 and A2 the $(q, r)^{th}$ element of $\boldsymbol{J}_{BP}$ for $(q, r) \in [\![1; Q]\!]^2$ is:*

$$J_{BP}[q, r](\Delta f_{q,r}) = $$
$$\frac{1}{N} \left(J_{Nyq}[q, r](\Delta f_{q,r}) \cdot H_{LP}(\Delta f_{q,r})\right) * $$
$$\sum_{k=-\infty}^{\infty} \delta(\Delta f_{q,r} - k f_p) \qquad (24)$$

where J_{Nyq} and J_{BP} are the Fisher information matrices for Nyquist sampling and Bandpass Sampling, $\Delta f_{q,r}$ is the distance in the frequency domain between $\tilde{x}_{\lambda_q}$ and $\tilde{x}_{\lambda_r}$, $H_{LP}(f)$ is the transfer function of an ideal low-pass filter with cut-off frequency $f_c = f_p/2$ and unit gain, $*$ the convolution operator, and $\delta(.)$ is the Dirac distribution. Notice that term $\sum_{k=-\infty}^{\infty} \delta(\Delta f_{q,r} - k f_p)$ corresponds to a f_p-spaced Dirac comb distribution, used as periodisation operator. The Cramér-Rao bound is then obtained through eq. (17).

Refer to Appendix C for a proof.

A. Discussions

The relationships established between Fisher matrices for compressive multiband, subsampling and Nyquist sampling allow to distinguish the influence on the estimation bound of the effect of subsampling from the effect of diversity creation between channels. Compared to inequality (19), an additional hypothesis on a multiband architecture and a disjoint subband model is made in the context of Spectrum Sensing to obtain (21) and (22). This assumption enables to describe the effect of the acquisition matrix on each frequency subband and thus enables to determine each term thanks to the deterministic nature of the sensing matrix in practice.
It is not surprising to see the Gramian appearing in Theorem 1. Due to dimensionality reduction, the Gramian $\boldsymbol{B}_{\boldsymbol{\Lambda}}^H \boldsymbol{B}_{\boldsymbol{\Lambda}}$ of a Compressive Sensing matrix $\boldsymbol{B}_{\boldsymbol{\Lambda}}$ is of maximum rank $M \leq N$. Hence neither norms nor orthogonality can be totally preserved by the projection through the sensing matrix $\boldsymbol{\Phi}$. The Gramian expresses this deformation of the scalar product by

the acquisition process. The interpretation for complex-valued x is less intuitive due to the effect of a complex Gramian on a complex signal, however the behaviour regarding parameters and the order of magnitude is similar as in the case where x is real-valued. Signals spread over multiple bands that do not satisfy Asm. A1 would induce non-trivial crossterms in the computations but similar global behaviour (dependency on the Gramian) could be expected. One fundamental but thin subtlety allowing to establish the first theorem lies in the fact that the sensing matrix of compressive multiband architectures can be expressed in the 'compact formalism' of eq. (10). Hence all frequencies of an analog subband are projected identically, even if they do not lie on the Nyquist grid with frequency resolution δ_f. This means that leakage provoked by the discretization of the analog disjoint model will depend on the coefficient of the sensing matrix $\boldsymbol{B}$ given by the original support index (cf proof Appendix C).

(24) means that $J_{BP}[q,r](\Delta f_{q,r})$ is f_p-periodic and matches with $J_{Nyq}[q,r]$ at baseband up to a noise folding factor N. The link between $\boldsymbol{J}_{BP}$ and $\boldsymbol{J}_{Nyq}$ is drawn under a light approximation (Asm. 2) that is more accurate as the acquisition time grows, because potential spectral leakage from off-frequency grid components is reduced.

V. APPLICATIONS TO COGNITIVE RADIO CANONICAL CASES

To illustrate the interest of this theorem for Cognitive Radio applications, we will focus on the canonical application case of amplitude and frequency estimation of two real tones. This admittedly simplified and well-known scenario is selected to highlight the influence of $\boldsymbol{B}$ on the expected estimation performance and mimics the detection of an interferer with background noise, a common issue for wideband compressive receivers, as mentionned in ([7], [19]). The noisy time domain input signal sampled at rate f_{Nyq} is given for $\tilde{n} \in [\![1; NL]\!]$ by:

$$x[\tilde{n}]^+ = A_1 \cos(2\pi f_1 \tilde{n} + \phi_1) + A_2 \cos(2\pi f_2 \tilde{n} + \phi_2) + w[\tilde{n}] \quad (25)$$

where $(f_1, f_2) \in [-0.5; 0.5]^2$ are normalized frequencies (by f_{Nyq}) and $w \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I}_{NL})$. $\boldsymbol{x}^+$ can be decomposed as:

$$\boldsymbol{x}^+ = \boldsymbol{E}\boldsymbol{a} + \boldsymbol{w} \quad (26)$$

with

$$\boldsymbol{E} = \begin{pmatrix} \cos(2\pi f_1 + \phi_1) & \cos(2\pi f_2 + \phi_2) \\ \vdots & \vdots \\ \cos(2\pi f_1 NL + \phi_1) & \cos(2\pi f_2 NL + \phi_2) \end{pmatrix}, \quad (27)$$

$$\text{where } \boldsymbol{E} \in \mathbb{R}^{NL \times 2} \text{ and } \boldsymbol{a}^T = (A_1, A_2).$$

Let us note $\Delta f = f_2 - f_1$, $\Delta\phi = \phi_2 - \phi_1$, $\Sigma f = f_2 + f_1$ and $\Sigma\phi = \phi_2 + \phi_1$. For illustrations, the MWC is chosen but the proposed Theorem 1 is valid for all multiband architectures (Bandpass, MWC, QAIC, PNUS, MRS); results can be easily extended by replacing the corresponding $\check{\boldsymbol{B}}$, as exposed in section II-B. The exhaustive Fisher matrix for Nyquist sampling of K real tones which is used as benchmark is given in Appendix D. For sake of clarity, disjoint estimation with

2 real tones is presented. Joint estimation or multitones is methodologically similar.

A. Amplitude estimation

The normalized frequencies of the useful signal f_1 and of a possible interferer f_2 are assumed to be known, and the parameters to estimate are $\boldsymbol{\theta} = (A_1, A_2)$. First, classical results for Fisher Information matrices are recalled for Nyquist sampling based amplitude estimation. A focus is put on the case of off-grid frequency, of practical interest and useful for the derivations in the next sections. Second, the role of Bandpass Sampling is studied. Finally these results are used to express the Cramér-Rao lower Bound for multiband based parametric estimation.

1) Nyquist samples: Using (17) (25) and according to Appendix D (79), the Fisher matrix for Nyquist Sampling is:

$$\boldsymbol{J}_{Nyq} \simeq \frac{NL}{2\sigma^2} \begin{pmatrix} 1 & J_{.,Nyq}[1,2] \\ J_{.,Nyq}[1,2] & 1 \end{pmatrix} \quad (28)$$

where $J_{.,Nyq}[1,2]$ is given by:

$$J_{.,Nyq}[1,2] = J_{1/2,Nyq}[1,2](\Delta f, \Delta\phi) + J_{1/2,Nyq}[1,2](\Sigma f, \Sigma\phi) \quad (29)$$

with:

$$J_{1/2,Nyq}[1,2](f,\phi) \simeq \quad (30)$$
$$\text{sinc}(\pi NL f) \cos(\phi + \pi f(NL+1))$$

where $\text{sinc}(x)$ stands for $\frac{\sin(x)}{x}$. The lower the cross-term coefficient of the Fisher matrix $J_.[1,2]$, the lower the determinant $\Delta = \frac{NL}{2\sigma^2}[(1 - J_.[1,2]^2)]$ of $\boldsymbol{J}_{Nyq}$ and the better the estimation. When $J[1,2] = 0$ the matrix is diagonal meaning that the ignorance of one amplitude does not degrade the estimation of the other. When $f_1 = f_2$ and $\phi_1 = \phi_2 \mod \pi$, or $f_1 = -f_2$ and $\phi_1 = -\phi_2 \mod \pi$, then $J[1,2] = J[1,1]$ and the matrix is singular, meaning that amplitude estimation by an unbiased estimator is not possible because foldings are indiscernible. Consider first, as often in the literature, the frequency on-grid case meaning that $\exists (k,k') \in \mathbb{N}^{*2}$ such that $\Delta f.f_{Nyq} = k\delta f$ and $\Sigma f.f_{Nyq} = k'\delta f$. Then $J_.[1,2] = 0$ and $\boldsymbol{J}_{Nyq} = \frac{NL}{2\sigma^2} \cdot \boldsymbol{I}_2$. There is no spectral leakage if T_{acq} is a multiple of $\frac{1}{f_2 - f_1}$ and $\frac{1}{f_2 + f_1}$. Furthermore the accuracy of the CRB depends only on the number of samples NL.

Consider now the more realistic off-grid case meaning that $\exists \, 0 < \eta, \eta' < 1$ such that:

$$\Delta f.f_{Nyq} = \left\lfloor \frac{(f_2 - f_1).f_{Nyq}}{\delta f} \right\rfloor \delta f + \eta \delta f$$

$$\Sigma f.f_{Nyq} = \left\lfloor \frac{(f_1 + f_2).f_{Nyq}}{\delta f} \right\rfloor \delta f + \eta' \delta f$$

where $\eta\delta f$ and $\eta'\delta f$ stands for the distance of $(f_2 - f_1)f_{Nyq}$ and $(f_1 + f_2)f_{Nyq}$ respectively to the grid.

Then eq. (31) transforms into:

$$J_.[1,2] = \frac{1}{NL} \Big[\frac{\sin(\pi(k+\eta))}{\sin(\pi\Delta f)} \cos(\Delta\phi + \pi\Delta f(NL+1)) \quad (31)$$
$$+ \frac{\sin(\pi(k+\eta'))}{\sin(\pi\Sigma f)} \cos(\Sigma\phi + \pi\Sigma f(NL+1)) \Big]$$

Hence as NL increases, $J_{.,Nyq}[1,2]$ approaches 0.

2) Bandpass Sampling: For Bandpass Sampling at normalized subsampling frequency $f_p = \frac{1}{N}$, let us define the frequencies folded at baseband with:

$$\bar{f}_1 = f_1 - f_p \left\lfloor \frac{f_1}{f_p} \right\rfloor, \bar{f}_2 = f_2 - f_p \left\lfloor \frac{f_2}{f_p} \right\rfloor \text{ and } \Delta\bar{f} = \bar{f}_1 - \bar{f}_2.$$

From (25) and (10), the output spectrum is:

$$\frac{\partial y[l]_{BP}}{\partial A_i} = \cos(2\pi \bar{f}_i N l + \phi_i), \text{ for } i = 1, 2. \quad (32)$$

Hence from (17):

$$\boldsymbol{J}_{BP} = \frac{L}{2\sigma^2} \begin{pmatrix} 1 & J_{.,BP}[1,2] \\ J_{.,BP}[1,2] & 1 \end{pmatrix} \quad (33)$$

where:

$$J_{.,BP}[1,2] = \frac{1}{L} \sum_{l=1}^{L} [\cos(2\pi\Delta\bar{f}Nl + \Delta\phi) + \cos(2\pi\Sigma\bar{f}Nl + \Sigma\phi)] \quad (34)$$

Similarly to the establishment of eq. (79) in Appendix D:

$$J_{.,BP}[1,2] = J_{1/2,BP}[1,2](N\Delta\bar{f}, \Delta\phi) + \quad (35)$$
$$J_{1/2,BP}[1,2](N\Sigma\bar{f}, \Sigma\phi)$$

where:

$$J_{1/2,BP}[1,2](f,\phi) \simeq \text{sinc}(\pi Lf) \cos(\phi + \pi f(L+1)) \quad (36)$$

3) Multiband CS: From Theorem 1, (22):

$$\boldsymbol{J}_{MB} = \Re \left[\boldsymbol{B}_\Lambda^H \boldsymbol{B}_\Lambda \right] \circ \boldsymbol{J}_{BP} \quad (37)$$

The results above on J_{MB} and J_{BP} will be discussed below.

4) Interpretation: Consider first the on-grid frequency case ($\exists (k, k') \in \mathbb{N}^{*2}$ s.t. $\Delta\bar{f}.f_{Nyq} = k\delta\bar{f}$ and $\Sigma\bar{f}.f_{Nyq} = k'\delta\bar{f}$). As we obtained for Nyquist samples, $J_{BP}[1,2] = J_{BP}[2,1] = J_{MB}[2,1] = J_{MB}[1,2] = 0$. The values taken by $J[1,1]$ and $J[2,2]$ are proportional to the number of samples, which is L for bandpass sampling, N times less than Nyquist sampling, for the same acquisition time. Hence when the two tones do not interfere with each others, Bandpass and Multiband compressive amplitude estimation compared to estimation from the Nyquist samples are exactly characterized by the respective compression ratios $1/N$ and M/N.

Consider now the off-grid case. The interference between two tones with Compressive Sensing is then influenced by the product of two terms. The first term is the real part of an element of the Gramian sensing matrix $B_\Lambda^H B_\Lambda[1,1]$ and would be one if frequencies belong to the same frequency subband. The second term, $J_{BP}[1,2]$, depends on the distance to the grid, as shown in eq. (35). For $\phi_1 = \phi_2$, one may notice that $f_1 = f_2$ for Nyquist and Multiband Compressive Sampling and that $f_1 = f_2 \mod (f_p)$ for Bandpass Sampling leads as expected to a singular J. The mismatches between the two tones model and the single tone observation makes the estimation problem singular. Results obtained in this Subsection are slightly more accurate than the Theorem 2 because Asm. A2 on bandlimited spectral representations of the sampled tones needs not be assumed.

5) Maximum Likelihood estimator: The Maximum-Likelihood estimator (MLE) is unbiased and efficient, meaning that its covariance tends to the Cramér-Rao Bound as the number of samples tends to infinity. Under noise gaussianity assumption the MLE in the non-compressed case is given from (26) by:

$$\hat{\boldsymbol{a}}_{\boldsymbol{Nyq}} = (\boldsymbol{E}^H \boldsymbol{E})^{-1} \boldsymbol{E}^H \boldsymbol{x}^+ \quad (38)$$

with $\boldsymbol{E}$ as in (27), and its covariance matrix by:

$$Cov(\hat{\boldsymbol{a}}_{Nyq}) = \sigma^2 (\boldsymbol{E}^H \boldsymbol{E})^{-1} \simeq \frac{2\sigma^2}{NL} \quad (39)$$

where the last approximation, holds for active frequencies different from 0 and 1/2, as detailed in [20]. For multiband architecture, we have $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{a}$ where

$$\boldsymbol{H} = \begin{pmatrix} \begin{pmatrix} \cos(2\pi\bar{f}_1 N + \phi_1) & \cos(2\pi\bar{f}_2 N + \phi_2) \\ \vdots & \vdots \\ \cos(2\pi\bar{f}_1 NL + \phi_1) & \cos(2\pi\bar{f}_2 NL + \phi_2) \end{pmatrix} \begin{pmatrix} b_{1,1} & 0 \\ 0 & b_{1,2} \end{pmatrix} \\ - - - - - - - - - - - - - - - \\ \vdots \\ - - - - - - - - - - - - - - - \\ \begin{pmatrix} \cos(2\pi\bar{f}_1 N + \phi_1) & \cos(2\pi\bar{f}_2 N + \phi_2) \\ \vdots & \vdots \\ \cos(2\pi\bar{f}_1 NL + \phi_1) & \cos(2\pi\bar{f}_2 NL + \phi_2) \end{pmatrix} \begin{pmatrix} b_{M,1} & 0 \\ 0 & b_{M,2} \end{pmatrix} \end{pmatrix} \quad (40)$$

is a matrix of size $ML \times 2$ and $b_{i,j}$ is the $(i,j)^{th}$ term of the multiband sensing matrix. If we assume that the compressed noise is white Gaussian with variance σ^2, the Maximum-Likelihood estimator coincides with the best linear estimator and is given by:

$$\hat{\boldsymbol{a}}_{MB} = (\boldsymbol{H}^H \boldsymbol{H})^{-1} \boldsymbol{H}^H \boldsymbol{y}^+ \quad (41)$$

and the covariance matrix of the estimation is given by:

$$Cov(\hat{\boldsymbol{a}}_{MB}) = \sigma^2 (\boldsymbol{H}^H \boldsymbol{H})^{-1} \quad (42)$$

For multiband architectures, from eq. (41), cross-terms of $\boldsymbol{J}$ are negligible toward diagonal terms as the number of samples tends to infinity as detailed in Appendix D, approaching the on-grid case:

$$Cov(\hat{\boldsymbol{a}}_{MB}) \approx \frac{2\sigma^2}{ML} \begin{pmatrix} \frac{1}{(\boldsymbol{B}_\Lambda^H \boldsymbol{B}_\Lambda)_{1,1}} & 0 \\ 0 & \frac{1}{(\boldsymbol{B}_\Lambda^H \boldsymbol{B}_\Lambda)_{2,2}} \end{pmatrix}$$
$$\approx \frac{N}{M} Cov(\hat{\boldsymbol{a}}_{Nyq}) \quad (43)$$

The empirical variance of the MLE obtained for both the Nyquist and Modulated Wideband Converter architecture is shown in Fig. 5. Results are also compared to the Cramér-Rao Bound from eq. (24). The empirical MLE variance indeed converges quickly to the Cramér-Rao Bound. Note also the expected $10 \log_{10}(M/N) = 4,04$ dB loss between the acquisition methods, which matches with the Signal-to-Noise degradation by the Compressive Sensing projection of 3 dB gain per doubling of the subsampling ratio $\frac{M}{N}$.

Now that mechanisms of estimation with compressive multiband architecture have been highlighted for amplitude estimation, another example, frequency estimation, will be tackled in the next Subsection. This time with a focus on the role of the Gramian of the sensing matrix.
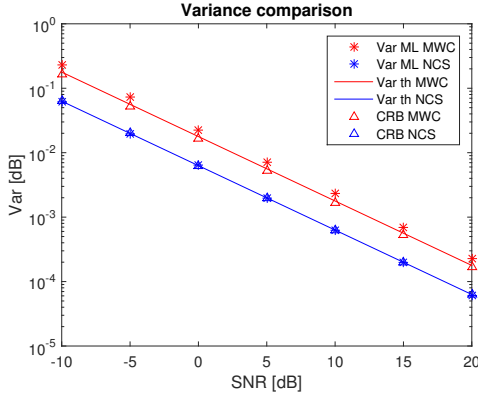
Fig. 5: Variance comparison for amplitude estimation of 2 real tones. Empirical (stars, from (38) and (41)) and Expected (triangle, from (39) and (42)) variance of the Maximum-Likelihood Estimator, analytical Cramér-Rao Bound from Theorem 2 (line), Blue=Nyquist sampling, Red=MWC with Random Bernoulli codes, $M = 50$, $N = 127$, 1000 trials, $f_1 = 1$ GHz, $f_2 = 3$ GHz, $f_{Nyq} = 10$ GHz, $A_1 = A_2 = 1$.

B. Frequency estimation

The second scenario mimics the detection of an interferer at any frequency with known amplitude (A_1, A_2) and phase (ϕ_1, ϕ_2) with background noise. The parameters to be estimated are the frequencies of the two tones $\boldsymbol{\theta} = (f_1, f_2)$. Although such a scenario is barely realistic, its purpose is to validate and illustrate the expressions found in the previous theorems and thus highlight the key role played by $\boldsymbol{B}$ in the multiband framework.

1) Nyquist samples: For $\boldsymbol{\theta} = (f_1, f_2)$, according to eq. (17) and (25) we have:

$$\boldsymbol{J}_{Nyq} \simeq \frac{2\pi^2}{\sigma^2} \begin{pmatrix} A_1^2 S_2(NL) & J_{.}[1,2] \\ J_{.}[1,2] & A_2^2 S_2(NL) \end{pmatrix}$$

where $S_2(NL) = \sum_{\tilde{n}=1}^{NL} \tilde{n}^2 = \frac{NL(NL+1)(NL+2)}{6}$ and from (81) with same notations as in (31):

$$J_{.,Nyq}[1,2] = A_1 A_2 \Re \left[e^{j\Delta\phi_{1,2}} \chi_2(\Delta f, NL) - e^{j\Sigma\phi_{1,2}} \chi_2(\Sigma f, NL) \right]$$
$$\text{where } \chi_2(f, NL) = \sum_{\tilde{n}=1}^{NL} \tilde{n}^2 e^{j2\pi f\tilde{n}}. \quad (44)$$

2) Bandpass: We have:

$$\frac{\partial y[l]_{BP}}{\partial f_i} = -2\pi N A_i l \sin(2\pi \bar{f}_i N l + \phi_i), \text{ for } i = 1, 2. \quad (45)$$

Therefore from (17) and similarly to the establishment of (81):

$$\boldsymbol{J}_{BP} \simeq \frac{2\pi^2}{\sigma^2} \begin{pmatrix} N^2 A_1^2 S_2(L) & J_{.,BP}[1,2] \\ J_{.,BP}[1,2] & N^2 A_2^2 S_2(L) \end{pmatrix}$$

where:

$$J_{.,BP}[1,2] = A_1 A_2 \Re \left[\sum_{l=1}^{L} l^2 e^{2j\pi\Delta\bar{f}Nl + \Delta\phi} \right]$$
$$= A_1 A_2 \Re \left[e^{j\Delta\bar{\phi}_{1,2}} \chi_2(\Delta\bar{f}N, L) - e^{j\Sigma\bar{\phi}_{1,2}} \chi_2(\Sigma\bar{f}N, L) \right]$$
$$(46)$$

And from Theorem 1 for multiband compressive architectures with sensing matrix $\boldsymbol{B}$ follows:

$$\boldsymbol{J}_{MB} = \Re\left[\boldsymbol{B}_\Lambda^H \boldsymbol{B}_\Lambda \right] \circ \boldsymbol{J}_{BP} \quad (47)$$

In accordance with [17], the variance in $O(\frac{\sigma^2}{(NL)^3})$ turns into $O(\frac{\sigma^2}{(NL)^2 ML})$ due to the dimension reduction. As the behaviour of eq. (46) is not straightforward, these expressions will be studied in simulation.

3) Simulation: The mixing codes of the MWC in the time domain are randomly generated from a Bernoulli distribution. For these three acquisition methods (Nyquist sampling, Bandpass Sampling and Multiband) the normalized coefficient $J_{norm}[1,2] = \frac{J[1,2]}{J[1,1]}$ is pictured in Fig. 6. Note that if it is equal to 0, there is no degradation of the estimation of one amplitude due to the lack of knowledge about the other amplitude. If it is equal to 1, it is impossible to raise the ambiguity between the two signals and no estimation can be made. The x-axis corresponds to the relative normalized frequency $\frac{f_2}{f_{Nyq}}$ and the term $\frac{f_1}{f_{Nyq}} = 0.22581$ is arbitrarily fixed. The value of $J_{1,2(norm)}(\Delta f)$ for Bandpass Sampling is
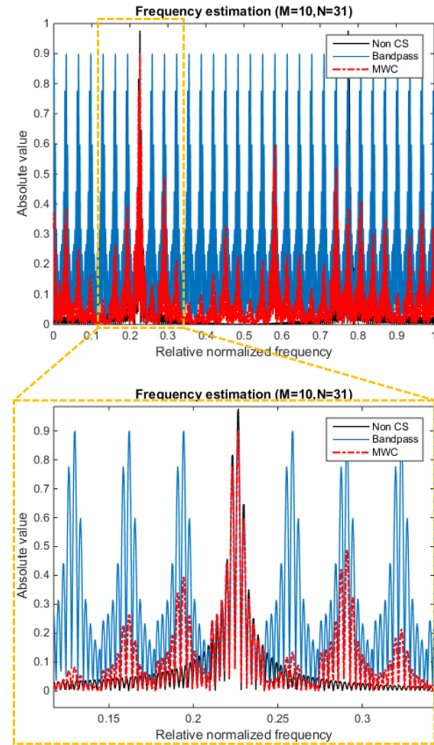

Fig. 6: Normalized off-diagonal coefficient of the Fisher matrix for frequency estimation $J_{norm}[1,2](\frac{f_2}{f_{Nyq}})$ (top=overall, bottom=zoom), $[M = 10, N = 31]$.

the f_p-periodic repetition of the value of $J_{1,2(norm)}(\Delta f)$ in $[-f_p/2; +f_p/2]$ for a Nyquist approach. For multiband Compressive Sensing, $J_{1,2(norm)}(\Delta f)$ is additionally weighted in each frequency subband. This weight is the crosscorrelation between the two support columns of the sensing matrix. It would be 1 for Bandpass Sampling. Hence for two frequencies f_1 and f_2 belonging to the same subband, they fold similarly and the Cramér-Rao Bound for Multiband CS is equal to
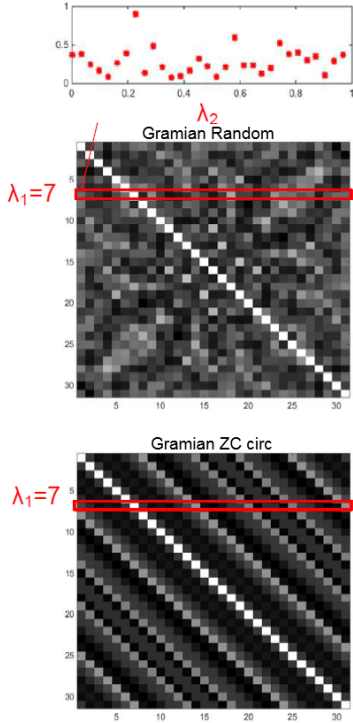
Fig. 7: Gramian of the Modulated Wideband Converter sensing matrix for different codes: Random Bernoulli (top) and Zadoff-Chu-based circulant (bottom) [N=31].

the Bandpass bound. However for two frequencies falling in different subbands an attenuation coefficient controlled by the sensing matrix appears. Note that the figure is not symmetric due to the influence of the second term depending on $\Sigma \bar{f}$, $\Re[-e^{j\Sigma\bar{\phi}}\chi_2(\Sigma\bar{f})]$ in eq. (46).

4) Influencing the sensing matrix Gramian: For sake of generality, $\boldsymbol{B}$ is frequently supposed to be random, however the sensing matrix depends on both the structure of the architecture and on a pseudo-random element (mixing codes, sample selection patterns, etc...). This opens the opportunity to design codes to improve the overall performance of the system. For example our findings may be used to take into account a non uniform input spectrum distribution. Element (i, j) of the Gramian is the crosscorrelation between columns i and j of the matrix. A simple example on how to influence the Gramian is given hereafter to yield a brief insight on the possibilities.

It is shown in [3] that the MWC sensing matrix may be expressed as $\boldsymbol{B} = \boldsymbol{A}\tilde{\boldsymbol{F}}\boldsymbol{D}$ where $\boldsymbol{A}$ is the time expression of the code, $\tilde{\boldsymbol{F}}$ is a reordered subset of the DFT matrix, $\boldsymbol{D}$ is a diagonal matrix accounting for pulse shaping. Fig. 7 presents the Gramian for a Modulated Wideband Converter, and two different codes studied in [21] for reconstruction. "Random" corresponds to a Bernoulli distribution, "ZC" corresponds to an example of structured codes with good crosscorrelation properties, Zadoff-Chu based circulant codes introduced in [21].

The weights of the coefficients for multiband architectures (red dots on the red curve) in Fig. 6 are given by the values

on the 7th row ($\boldsymbol{\Lambda}(f_1) = 7$) of the Gramian in Fig. 7. For the two different structures of codes, the Gramian presents two specific patterns. For "Random" codes, values off the diagonal appear uniformly distributed whereas the circulant character of code (b) implies equal-valued diagonal lines. Hence by changing the shift of the circulant matrix, it is possible to choose the circular permutation with the most favorable coefficient between λ_1 and λ_2. More sophisticated tunnings, by optimizing the choice of M rows among N for example, are also possible. Note that a more thorough analysis available in [21] shows that the coherence of code (b) is also lower. Thus it is possible to adapt the structure of the mixing code in the MWC, and the choice of the pseudo-random element in the general case to promote non-uniformly distributed performance according to the frequency subband index. This result shows that it is possible to control the Cramér-Rao Bound and hence the frequency estimation precision thanks to the crosscorrelation properties of the acquisition matrix.

With the random codes MWC given in example, $\Re\left[\boldsymbol{B}_{\boldsymbol{\Lambda}}^H \boldsymbol{B}_{\boldsymbol{\Lambda}}[4,7]\right] = 0.9$, so frequency estimation of a single tone signal is as accurate if there is an interferer in the 4th subband with -10 dB Signal-to-Interference ratio (SIR), as if the interferer is in the 7th subband (same as the signal, $\boldsymbol{\Lambda}(f_1) = 7$), same bin, with 0 dB SIR. Given the orders of magnitude obtained, it is possible to handle even strong blockers, if the sensing matrix is adequately chosen.

VI. CONCLUSION

A. Summary

In this article a study of the theoretical performances of spectral parameter estimation was led based on the Cramér-Rao lower Bound. First a framework of compressive radiofrequency architectures was defined. For multiband compressive architectures and disjoint frequency subband models a new theorem which provides a benchmark toward Nyquist sampling has been established. Relationships between Fisher matrices are exposed: first between Compressive Sampling and subsampling, then between subsampling and Nyquist sampling. The role of the Gramian of the sensing matrix restricted to the support is highlighted. Illustrations were given for concrete scenarios of Cognitive Radio, namely the estimation of amplitude and frequency of an out-of-band interferer. For amplitude estimation, the on-grid case yields a diagonal Fisher matrix, and bounds are exactly described by the compression ratio. Otherwise performance depends on the crossterms of J, and therefore on correlations between the columns of the sensing matrix that belong to the support. Since the precision of frequency and amplitude estimation can be controlled for each pair of subbands through the crosscorrelation properties of the sensing matrix, a simple way to optimize performances is pointed out. A sketch is given on how to adapt the sensing matrix to priors on the spectrum distribution by choosing adequate codes. This represents a potential workaround for the crucial interferer folding issue.

B. Discussions

The presented result offers a bound on each term of the Fisher matrix from samples acquired with a multiband

compressive radiofrequency receiver. As it is adapted to each term, it is more accurate than [10]. This is achieved thanks to the deterministic nature of the sensing matrix in practice and assuming a specific signal model. It is interesting to note that since reconstruction performance depends on coherence and norm preservation, metrics which are closely linked to the Gramian of $\boldsymbol{B}$, a good matrix for feature extraction is a good matrix for reconstruction and vice versa. This is not surprising as in this educational scenario the signal is accurately described by the set of parameters.

The suggested adaptive approach is flexible: compared for example to the interference cancellation proposed in [7], new codes need not be generated at every interferer detection.

The proposed framework is relatively general as it concerns all types of parameters and all compressive multiband architectures, whether for real- or complex-valued signals (after frequency downconversion). An interesting extension of the signal model would be to go past the assumption of bandlimited representation of the samples of the signal (Asm. A2).

C. Perspectives

Cramér-Rao Bounds expression for other wideband RF receivers and wider class of signals ought to be exposed as well. Also average and worst-case behaviours of the cross-terms depending on dephasing and distance to the grid could be characterized more thoroughly. A very interesting question to investigate remains about the possibility to derive biased estimators with lower variance.

APPENDIX A
MWC SENSING MATRIX

For sake of completeness, the expression of $\check{\boldsymbol{B}}$ for the MWC is established. The mth mixing code admits following Fourier decomposition:

$$p_m(t) = \sum_{n'=-\infty}^{+\infty} \gamma_{m,n'} e^{2i\pi t \frac{n'}{T_p}} \tag{48}$$

Its restriction to a period T_p is given by:

$$p_{m,[T_p]}(t) = \sum_{q=0}^{N-1} \alpha_m(q) g(t - qT_c) \tag{49}$$

where $\boldsymbol{\alpha}_m$ corresponds to the N code elements and g is a shaping function, by default a rectangular window of length T_c. Thus the Fourier Transform of the signal after mixing:

$$FT[x(t)p_m(t)] = \tilde{x}(f) * \tilde{p}_m(f) = \sum_{n'=-\infty}^{+\infty} \gamma_{m,n'} \tilde{x}(f - \frac{n'}{T_p})$$

$$= \sum_{n'=-\infty}^{+\infty} \gamma_{m,-n'} \tilde{x}(f + \frac{n'}{T_p}) \tag{50}$$

Yet:

$$\gamma_{m,n'} = \frac{1}{T_p} \tilde{p}_{m,[T_p]}[\frac{n'}{T_p}] = \frac{1}{T_p} \tilde{g}(\frac{n'}{T_p}) DFT(\boldsymbol{\alpha}_m)[n'] \tag{51}$$

Note that $\tilde{x}(f)$ can be rewritten as a sum of shifted support limited functions:

$$\tilde{x}(f) = \sum_{k=-\lfloor N/2 \rfloor}^{\lfloor N/2 \rfloor} \tilde{\beta}_k(f - kf_p) \tag{52}$$

where $\tilde{\beta}_k(f) = 0$ if $f \notin [-\frac{f_p}{2}; \frac{f_p}{2}]$.
Hence eq. (50) can be rewritten:

$$FT[x(t)p_m(t)] =$$
$$\sum_{k=-\lfloor \frac{N}{2} \rfloor}^{\lfloor \frac{N}{2} \rfloor} \sum_{n'=-\lfloor \frac{N}{2} \rfloor}^{\lfloor \frac{N}{2} \rfloor} \gamma_{m,-n'} \tilde{\beta}_k(f - (k - n')f_p) \tag{53}$$

After filtering, since $\tilde{\beta}_k$ are support limited, only $k = n'$ is kept at baseband:

$$\tilde{y}_m(f) = FT[x(t)p_m(t)]_{BB} = \sum_{n'=-\lfloor N/2 \rfloor}^{\lfloor N/2 \rfloor} \gamma_{m,-n'} \tilde{\beta}_{n'}(f) \tag{54}$$

For a given acquisition time, suppose that the spectrum does not have very quick fluctuations, we may write the decomposition of the continuous $\tilde{\beta}_k$ on the l discrete frequency bins by averaging for $k \in [\![-\lfloor N/2 \rfloor; \lfloor N/2 \rfloor]\!]$:

$$\tilde{\beta}_k(f) = \sum_{l=1}^{L} \tilde{\beta}_{k,l}(f - (l-1)\delta f) \tag{55}$$

Hence eq. (54) can be rewritten as:

$$\tilde{y}_m(f) = \sum_{n'=-\lfloor N/2 \rfloor}^{\lfloor N/2 \rfloor} \gamma_{m,-n'} \tilde{\beta}_{n'}(f) \tag{56}$$

$$= \sum_{l=1}^{L} \sum_{n'=-\lfloor N/2 \rfloor}^{\lfloor N/2 \rfloor} \gamma_{m,-n'} \tilde{\beta}_{n',l}(f) \tag{57}$$

where $\tilde{\beta}_{n',l}(f)$ represents the $\tilde{n}^{th} = (n-1)L + l^{th}$ bin of the input spectrum. This is because n' spans $[\![-\lfloor N/2 \rfloor; \lfloor N/2 \rfloor]\!]$ but index of the sensing matrix $n \in [\![1; N]\!]$. In other words:

$$\tilde{y}_m[l] = \tilde{y}_m \left(f = (l-1)\delta f - \frac{f_p}{2} \right)$$

$$= \sum_{n=1}^{N} b_{m,n} \tilde{x} \left((l-1)\delta f + (n - \lfloor \frac{N}{2} \rfloor - \frac{3}{2}).f_p \right) \tag{58}$$

for $l \in [\![1; L]\!]$. This demonstrates that the acquisition matrix $\check{\boldsymbol{B}}$ in eq. (7) is indeed given by $b_{m,n} = \gamma_{m,-(n-\lfloor \frac{N}{2} \rfloor -1)}$ with $\gamma_{m,n}$ defined by eq. (48) as the Fourier coefficients of the shaped mixing codes. ∎

APPENDIX B
PROOF OF THEOREM 1

Proof: First, (22) is derived, then (21) as a simplification. According to (17),(1) the $(q,r)^{th}$ element of the Fisher matrix from the compressed samples of a multiband architecture (MB) $\boldsymbol{J}_{MB}$ is given by:

$$J_{MB,[q,r]} = \frac{2}{\sigma^2} \Re \left[\sum_{\tilde{m}=1}^{ML} \frac{\partial y^*[\tilde{m}]}{\partial \theta_q} \frac{\partial y[\tilde{m}]}{\partial \theta_r} \right] \tag{59}$$

Due to orthonormality and independence toward parameters of the Discrete Fourier Transform matrix $\boldsymbol{F}$, it is also true in the spectral domain:

$$J_{MB,[q,r]} = \frac{2}{\sigma^2} \Re \left[\sum_{\tilde{m}=1}^{ML} \frac{\partial \tilde{y}^*[\tilde{m}]_{MB}}{\partial \theta_q} \frac{\partial \tilde{y}[\tilde{m}]_{MB}}{\partial \theta_r} \right] \quad (60)$$

For Bandpass Sampling, the output spectrum is:

$$\tilde{y}_{BP}(f) = \sum_{k=-\infty}^{+\infty} \tilde{x}(f - k f_p) \quad (61)$$

The $(q,r)^{th}$ element of the Fisher matrix is given by:

$$J_{BP}[q,r] = \frac{2}{\sigma^2} \Re \left[\sum_{l=1}^{L} \frac{\partial \tilde{y}^*[l]_{BP}}{\partial \theta_q} \frac{\partial \tilde{y}[l]_{BP}}{\partial \theta_r} \right] \quad (62)$$

where from the discretization of (61) the sample derivatives are given by:

$$\frac{\partial \tilde{y}[l]_{BP}}{\partial \theta_q} = \sum_{n=1}^{N} \frac{\partial \tilde{x}[L(n-1)+l]}{\partial \theta_q} = \frac{\partial \tilde{x}_q[L(\lambda_q - 1)+l]}{\partial \theta_q} \quad (63)$$

where the second equality is due to the subband model (Asm. A1). Whereas for Multiband Compressive Sensing, the samples derivatives are given, from (58), by:

$$\frac{\partial \tilde{y}_m[l]_{MB}}{\partial \theta_q} = \sum_{n=1}^{N} b_{m,n} \frac{\partial \tilde{x}[L(n-1)+l]}{\partial \theta_q}$$
$$= b_{m,\lambda_q} \frac{\partial \tilde{x}_q[L(\lambda_q-1)+l]}{\partial \theta_q} = b_{m,\lambda_q} \frac{\partial \tilde{y}[l]_{BP}}{\partial \theta_q} \quad (64)$$

where the second equality is due to (Asm. 1) and the third by recognizing (63). Hence:

$$\sum_{m=1}^{M} \sum_{l=1}^{L} \frac{\partial \tilde{y}_m^*[l]_{MB}}{\partial \theta_q} \frac{\partial \tilde{y}_m[l]_{MB}}{\partial \theta_r}$$
$$= \sum_{m=1}^{M} b_{m,\lambda_q}^* b_{m,\lambda_r} \sum_{l=1}^{L} \frac{\partial \tilde{y}^*[l]_{BP}}{\partial \theta_q} \frac{\partial \tilde{y}[l]_{BP}}{\partial \theta_r} \quad (65)$$

Yet:

$$\Re \left[b_{m,\lambda_q}^* b_{m,\lambda_r} \frac{\partial \tilde{y}^*[l]_{BP}}{\partial \theta_q} \frac{\partial \tilde{y}[l]_{BP}}{\partial \theta_r} \right] =$$
$$\quad (66)$$
$$\Re \left[b_{m,\lambda_q}^* b_{m,\lambda_r} \right] . \Re \left[\frac{\partial \tilde{y}^*[l]_{BP}}{\partial \theta_q} \frac{\partial \tilde{y}[l]_{BP}}{\partial \theta_r} \right] -$$
$$\Im \left[b_{m,\lambda_q}^* b_{m,\lambda_r} \right] . \Im \left[\frac{\partial \tilde{y}^*[l]_{BP}}{\partial \theta_q} \frac{\partial \tilde{y}[l]_{BP}}{\partial \theta_r} \right]$$

We recognize that:

$$\Im \left[\frac{\partial \tilde{y}^*[l]_{BP}}{\partial \theta_q} \frac{\partial \tilde{y}[l]_{BP}}{\partial \theta_r} \right] = -\Re \left[j \frac{\partial \tilde{y}^*[l]_{BP}}{\partial \theta_q} \frac{\partial \tilde{y}[l]_{BP}}{\partial \theta_r} \right] . \quad (67)$$

However $j \frac{\partial \tilde{y}^*[l]_{BP}}{\partial \theta_q} = \frac{\partial e^{j\frac{\pi}{2}} \tilde{y}^*[l]_{BP}}{\partial \theta_q}$, it corresponds to a simple additional dephasing of $\frac{\pi}{2}$ between the two subsignals. So according to (60) and (62), it is possible to rewrite matrixwise for $(q,r) \in [\![1;Q]\!]^2$:

$$\boldsymbol{J}_{MB} = \Re \left[\boldsymbol{B}_\Lambda^H \boldsymbol{B}_\Lambda \right] \circ \boldsymbol{J}_{BP} + \Im \left[\boldsymbol{B}_\Lambda^H \boldsymbol{B}_\Lambda \right] \circ \check{\boldsymbol{J}}_{BP} \quad (68)$$

Similarly but simpler in the real-valued case, by using eq. (17), $\boldsymbol{J}$ can be rewritten matrixwise for $(q,r) \in [\![1;Q]\!]^2$ into:

$$\boldsymbol{J}_{MB} = \Re \left[\boldsymbol{B}_\Lambda^H \boldsymbol{B}_\Lambda \right] \circ \boldsymbol{J}_{BP} \quad (69)$$

∎

APPENDIX C
PROOF OF THEOREM 2

Proof: Let us remind that $\Delta f_{q,r} = f_q - f_r$. We will first note that $J_{BP}[q,r](\Delta f_{q,r})$ is f_p-periodic and then demonstrate that $J_{BP}[q,r](\Delta f_{q,r}) = \frac{1}{N} J_{Nyq}[q,r](\Delta f_{q,r})$ for $\Delta f_{q,r} \in [-\frac{f_s}{2}; \frac{f_s}{2}]$, i.e. at lowpass.
First notice from (61) that $\tilde{y}$ is f_p-periodic, hence it follows from (62) that $J_{BP}[q,r](\Delta f_{q,r})$ is also f_p-periodic.
Concerning the second part, the Fisher information matrix for a Nyquist-rate acquisition is given by:

$$J_{Nyq}[q,r] = \quad (70)$$
$$\frac{2}{\sigma^2} \Re \left[\sum_{l=1}^{L} \sum_{n=1}^{N} \frac{\partial \tilde{x}^*[(n-1)L+l]}{\partial \theta_q} \frac{\partial \tilde{x}[(n-1)L+l]}{\partial \theta_r} \right]$$

According to Asm. A2, for a given $l \in [\![1;L]\!]$, the derivative:

$$\frac{\partial \tilde{x}_q[(n-1)L+l]}{\partial \theta_q}, \forall q \in [\![1;Q]\!] \quad (71)$$

for $n \in [\![1;N]\!]$ has at most one non-zero term, occurring for $n = \lambda_q$. Therefore for $\Delta f_{q,r} \in [-\frac{f_p}{2}; \frac{f_p}{2}]$ it is possible to simplify the sum over n leading to:

$$J_{Nyq}[q,r] = \quad (72)$$
$$\frac{2}{\sigma^2} \Re \left[\sum_{l=1}^{L} \frac{\partial \tilde{x}_q^*[(\lambda_q-1)L+l]}{\partial \theta_q} \frac{\partial \tilde{x}_r[(\lambda_r-1)L+l]}{\partial \theta_r} \right]$$

where λ_q and λ_r are equal, or consecutive indices (in case of leakage from the adjacent subband). Hence by recognizing (62)(63), we have $J_{Nyq}[q,r] = N J_{BP}[q,r]$ for $\Delta f_{q,r} \in [-\frac{f_p}{2}; \frac{f_p}{2}]$.
∎

APPENDIX D
CRB FOR NYQUIST SAMPLING OF K-TONE SIGNALS

Proof: Consider a K-tone signal, generalization of (25); the parameter vector is $\boldsymbol{\theta} = [A_1...A_K, \Phi_1...\Phi_K, f_1...f_K]$. For $k \in [\![1;K]\!]$:

$$\frac{\partial \tilde{x}[\tilde{n}]}{\partial A_k} = \cos(2\pi f_k \tilde{n} + \phi_k) \quad (73)$$

$$\frac{\partial \tilde{x}[\tilde{n}]}{\partial \phi_k} = -A_k \sin(2\pi f_k \tilde{n} + \phi_k) \quad (74)$$

$$\frac{\partial \tilde{x}[\tilde{n}]}{\partial f_k} = -A_k 2\pi \tilde{n} \sin(2\pi f_k \tilde{n} + \phi_k) \quad (75)$$

Then for Nyquist sampling, the coefficients of the Fisher matrix $\boldsymbol{J}_{Nyq}$ are given as follows.
For same parameters, $(k,k') \in [\![1;K]\!]^2$:

$$J[A_k, A_{k'}] = \frac{1}{\sigma^2} \sum_{\tilde{n}=1}^{NL} \cos(2\pi f_k \tilde{n} + \phi_k) \cos(2\pi f_{k'} \tilde{n} + \phi_{k'}) \quad (76)$$

Recall that $\cos(a)\cos(b) = \frac{1}{2}(\cos(a+b)+\cos(a-b))$ for $(a,b) \in \mathbb{R}$. Consider then the real part of a sum of geometrical terms with common ratio $e^{2j\pi f}$ given by ($f \neq 0$):

$$\sum_{\tilde{n}=1}^{NL} e^{2j\pi f\tilde{n}+\phi} = e^{\phi+2j\pi f}\frac{1-e^{2j\pi NLf}}{1-e^{2j\pi f}} \qquad (77)$$

Then from half angle factorization ($f \neq \{0; 0.5\}$):

$$e^{\phi+2j\pi f}\frac{1-e^{2j\pi NLf}}{1-e^{2j\pi f}} = e^{\phi+j\pi f(NL+1)}\frac{\sin(2\pi NLf)}{\sin(2\pi f)} \qquad (78)$$

Replacing in eq. (76) with the real part in eq. (78) yields:

$$J[A_k, A_{k'}] = \frac{1}{2\sigma^2}[$$
$$\frac{sin(NL(\Delta f)_{k,k'})}{sin((\Delta f)_{k,k'})}cos((\Delta\Phi)_{k,k'} + \pi(\Delta f)_{k,k'}(NL+1)) +$$
$$\frac{sin(NL(\Sigma f)_{k,k'})}{sin((\Sigma f)_{k,k'})}cos(\Sigma\Phi_{k,k'} + \pi(\Sigma f)_{k,k'}(NL+1))]$$

$$\simeq \frac{NL}{2\sigma^2}[sinc(NL(\Delta f)_{k,k'})cos((\Delta\Phi)_{k,k'} + \pi(\Delta f)_{k,k'}(NL+1)) +$$
$$sinc(NL(\Sigma f)_{k,k'})cos(\Sigma\Phi_{k,k'} + \pi(\Sigma f)_{k,k'}(NL+1))] \qquad (79)$$

Similarly by using $\sin(a)\sin(b) = \frac{1}{2}(-\cos(a+b)+\cos(a-b))$ for $(a, b) \in \mathbb{R}^2$:

$$J[\Phi_k, \Phi_{k'}] \simeq \frac{A_k A_{k'} NL}{2\sigma^2}[\qquad (80)$$
$$[sinc(NL(\Delta f)_{k,k'})\cos((\Delta\Phi)_{k,k'} + \pi(\Delta f)_{k,k'}(NL+1)) -$$
$$sinc(NL(\Sigma f)_{k,k'})\cos((\Sigma\Phi)_{k,k'}) + \pi(\Sigma f)_{k,k'}(NL+1))]$$

$$J[f_k, f_{k'}] =$$
$$\frac{2\pi^2}{\sigma^2}A_k A_{k'}\Re\left[e^{j(\Delta\Phi)_{k,k'}}\chi_2(\Delta f, NL) - e^{j(\Sigma\Phi)_{k,k'}}\chi_2(\Sigma f, NL)\right]$$
$$\text{where } \chi_2(f, NL) = \sum_{\tilde{n}=1}^{NL}\tilde{n}^2 e^{j2\pi f\tilde{n}}. \qquad (81)$$

For diagonal terms, $k \in [\![1; K]\!]$:

$$J[A_k, A_k] = \sum_{\tilde{n}=1}^{NL}\left(\frac{1}{2} + \frac{\cos(4\pi f_k + 2\phi_k)}{2}\right) \simeq \frac{NL}{2\sigma^2} \qquad (82)$$

$$J[\Phi_k, \Phi_k] \simeq \frac{NLA_k^2}{2\sigma^2} \qquad (83)$$

$$J[f_k, f_k] \simeq \frac{2\pi^2 A_k^2}{\sigma^2}\sum_{n=1}^{NL}n^2$$

$$= \frac{2\pi^2 A_k^2}{\sigma^2}\frac{NL(NL+1)(NL+2)}{6} \qquad (84)$$

For other terms ($k \neq k'$):

$$J[A_k, \Phi_{k'}] = \qquad (85)$$
$$-A_{k'}\frac{NL}{\sigma^2}\sum_{\tilde{n}=1}^{NL}\sin(2\pi f_k\tilde{n} + \phi_k)\cos(2\pi f_{k'}\tilde{n} + \phi_{k'})$$

Using $\sin(a)\cos(b) = \frac{1}{2}(\sin(a+b)+\sin(a-b))$ for $(a,b) \in \mathbb{R}$:

$$J[A_k, \Phi_{k'}] \simeq \frac{-A_{k'}NL}{2\sigma^2}[\qquad (86)$$
$$sinc(NL(\Delta f)_{k,k'})\sin((\Delta\Phi)_{k,k'} + \pi(\Delta f)_{k,k'}(NL+1)) +$$
$$sinc(NL(\Sigma f)_{k,k'})\sin((\Sigma\Phi)_{k,k'} + \pi(\Sigma f)_{k,k'}(NL+1))]$$

$$J[A_k, f_{k'}] = \frac{2\pi^2}{\sigma^2}A_{k'}\Im[e^{j(\Delta\Phi)_{k,k'}}\chi_2(\Delta f, NL) +$$
$$e^{j(\Sigma\Phi)_{k,k'}}\chi_2(\Sigma f, NL)] \qquad (87)$$

$$J[\Phi_k, f_{k'}] = \frac{-2\pi}{\sigma^2}A_k A_{k'}.$$
$$\Re\left[e^{j(\Delta\Phi)_{k,k'}}\chi_1(\Delta f, NL) - e^{j(\Sigma\Phi)_{k,k'}}\chi_1(\Sigma f, NL)\right] \qquad (88)$$

where $\chi_1(f, NL) = \sum_{\tilde{n}=1}^{NL}ne^{j2\pi f\tilde{n}}$.
For $k = k'$:

$$J[A_k, \Phi_k] \simeq$$
$$\frac{NL}{2\sigma^2}\text{sinc}(2NLf_k)\sin(2\Phi_k + 2\pi f_k(NL+1)) \qquad (89)$$

$$J[A_k, f_k] = \frac{2\pi^2}{\sigma^2}A_k\Im\left[e^{j2\Phi_k}\chi_2(2f_k, NL)\right]$$
$$J[\Phi_k, f_k] = \frac{2\pi}{\sigma^2}A_k^2\frac{NL(NL+1)}{2} \qquad (90)$$

Note that eq. (82-84) and (89-90) match eq. (3.41) in [16]. It is possible to compute further χ_1 and χ_2, either by simple but tedious calculations consisting in deriving twice $\sum_{n=1}^{N}e^{j2\pi fn}$ or by recognizing the Fourier Transform of a ramp function and a square function. ∎

APPENDIX E

NOTATION TABLE

Matrices and vectors will be written with uppercase and lowercase boldface letters respectively.

Notations	Signification
$.^T$	Transpose
$.^*$	Conjugate
$.^H$	Hermitian conjugate
$.^{-1}$	Inverse
$*$	Convolution
$\circ$	Hadamard product
$Tr(.)$	Trace operator
$\hat{}$	Estimation
$\lfloor . \rfloor$	Floor operator
$\mathrm{mod}\,(.)$	Modulo operator
$< .,. >$	Scalar product
$[\![\ldots ; \ldots]\!]$	Interval of integers
$\Re(.)$	Real part
$\Im(.)$	Imaginary part
$\lvert . \rvert$	Module
$x(t)$	Analog time-domain input signal
$y(t)$	Analog time-domain output signal
$\boldsymbol{x}[.]$	Discrete time-domain input vector
$\tilde{x}(f)$	Fourier Transform of $x(t)$
$\tilde{\boldsymbol{x}}[.]$	Discrete Fourier Transform (DFT) of $x(t)$
$x^+(t)$	Noisy analog time-domain input signal
$x_\sharp$	Reordered $N \times L$ input signal for the compact multiband formalism
$y_\sharp$	Reordered $M \times L$ output signal for the compact multiband formalism
f_p	Width of a subband in the signal model
T_p	Period of the mixing code
f_s	2x the filter cut-off frequency
T_s	Sampling period
δf	Frequency resolution
T_{Nyq}	Time resolution of the input signal
T_{acq}	Acquisition length such that there are NL input samples
f_{Nyq}	Nyquist sampling rate of $x(t)$
T_c	Chip period of the code in the MWC
K	Degree of sparsity of $x(t)$
L	Number of bins of width δf in a subband f_p or ratio between acquisition time and period of the codes
M	Number of branches, typically
N	Number of subbands
$\mu = ML$	Output dimension
$\nu = NL$	Input dimension
n	Subband index
l	Bin index within a subband
$\tilde{n}$	Global frequency index $\tilde{n} = (n-1)L + l$
Q	Length of the parameter vector
$\boldsymbol{\Lambda}$	Support
$\Lambda(.)$	Support index indicator function
λ_q	Support index of the qth parameter
$\tilde{x}_q$	Subsignal influenced by the qth parameter
$\Delta f_{q,r}$	$-f_q + f_r$
$\Sigma f_{q,r}$	$f_q + f_r$
$\Delta\phi_{q,r}$	$-\phi_q + \phi_r$
$\Sigma\phi_{q,r}$	$\phi_q + \phi_r$
$\bar{f}$	f folded at baseband by subsampling
β_k	Support-limited function representing the kth subband
δ_K	RIP constant of sparsity K
$\eta_{1,2}$	Offset to the frequency grid of Δf, in freq. resolution < 1
$\eta'_{1,2}$	Offset to the frequency grid of Σf, in freq. resolution < 1
$\gamma_{m,n}$	Fourier coefficients of the shapped mixing codes in the MWC
σ	White noise variance
$\boldsymbol{\theta}$	Vector of estimation parameter
$\boldsymbol{A}_{i,.}$	i^{th} row of $\boldsymbol{A}$
$\boldsymbol{A}_{.,i}$	i^{th} column of $\boldsymbol{A}$
$\boldsymbol{A}$	Matrix expressing, for the MWC the mixing codes in the time domain, and for the NUWBS the operations before time subsampling
$\boldsymbol{B}$	Multiband architecture matrix, compact formalism
$\check{\boldsymbol{B}}$	Multiband architecture matrix, canonical formalism
$\boldsymbol{D}$	Diagonal matrix accounting for pulse shaping
$\boldsymbol{E}$	Matrix model of the dual-tone signal at Nyquist time
$\boldsymbol{F}$	Discrete Fourier Transform matrix
$\tilde{\boldsymbol{F}}$	Reordered Discrete Fourier Transform matrix
$\boldsymbol{H}$	Matrix model of the dual-tone signal with CS
$\boldsymbol{J}$	Fisher information matrix
$\boldsymbol{Z}$	Sensing matrix applied on the sparsity domain $\boldsymbol{Z} = \boldsymbol{\Phi}\boldsymbol{\Psi}$
$\boldsymbol{\Gamma}$	Covariance matrix
$\boldsymbol{\Phi}$	Acquisition/Sensing matrix describing the Compressive Sensing acquisition process on the time signal
$\boldsymbol{\Psi}$	Matrix corresponding to the basis change to the sparsity domain (Sparsifying matrix)
$\boldsymbol{\Phi}\boldsymbol{\Psi}$	Acquisition/Sensing matrix describing the Compressive Sensing acquisition process on the sparse representation
$\mathbb{N}^*$	Set of strictly positive integers

REFERENCES

[1] E. J. Candes *et al.*, "Compressive Sampling," *Proc. of the International Congress of Mathematicians*, 2006.

[2] S. K. Sharma, E. Lagunas, S. Chatzinotas, and B. Ottersten, "Application of Compressive Sensing in Cognitive Radio Communications: A Survey," *IEEE Communications Surveys and Tutorials*, vol. 18, pp. 1838–1860, 2016.

[3] M. Mishali and Y. C. Eldar, "From Theory to Practice: Sub-Nyquist Sampling of Sparse Wideband Analog Signals," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, pp. 375–391, 2010.

[4] L. Yuan-Pei and P. P. Vaidyanathan, "Periodically nonuniform sampling of bandpass signals," *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Trans.*, vol. 45, no. 3, pp. 340–351, 1998.

[5] R. Maleh, G. L. Fudge, F. A. Boyle, and P. E. Pace, "Analog-to-Information and the Nyquist Folding Receiver," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, no. 3, pp. 564–578, 2012.

[6] M. Pelissier and C. Studer, "Non-uniform Wavelet Sampling for RF Analog-to-Information Conversion," *Circuits and Systems I: Regular Papers, IEEE Trans.*, vol. 65, no. 2, pp. 471–484, 2017.

[7] D. Adams, Y. Eldar, and B. Murmann, "A mixer frontend for a four-channel Modulated Wideband Converter with 62 db blocker rejection," in *IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, 2016, pp. 286–289.

[8] T. Haque, M. Bajor, Y. Zhang, J. Zhu, Z. A. Jacobs, R. B. Kettlewell, J. Wright, and P. R. Kinget, "A Reconfigurable Architecture Using a Flexible LO Modulator to Unify High-Sensitivity Signal Reception and Compressed-Sampling Wideband Signal Detection," *Solid-State Circuits, IEEE Journal of*, vol. 53, no. 6, pp. 1577–1591, 2018.

[9] Z. Ben-Haim and Y. C. Eldar, "The Cramer-Rao Bound for Estimating a Sparse Parameter Vector," *Signal Processing, IEEE Trans.*, vol. 58, no. 6, pp. 3384–3389, 2010.

[10] D. Ramasamy, S. Venkateswaran, and U. Madhow, "Compressive estimation in AWGN: General observations and a case study," in *Proc. Signals, Systems and Computers (ASILOMAR)*, 2012, Conference Proceedings, pp. 953–957.

[11] E. Arias-Castro and Y. C. Eldar, "Noise folding in compressed sensing," *IEEE Signal Processing Letters*, vol. 18, no. 8, pp. 478–481, 2011.

[12] R. G. Vaughan, N. L. Scott, and D. R. White, "The Theory of Bandpass Sampling," *Signal Processing, IEEE Trans.*, vol. 39, no. 9, pp. 1973–1984, 1991.

[13] S. Hongjian, C. Wei-Yu, J. Jing, A. Nallanathan, and H. V. Poor, "Wideband spectrum sensing with sub-nyquist sampling in cognitive radios," *Signal Processing, IEEE Trans.*, vol. 60, no. 11, pp. 6068–6073, 2012.

[14] T. Haque, R. T. Yazicigil, K. J. L. Pan, J. Wright, and P. R. Kinget, "Theory and design of a quadrature analog-to-information converter for energy-efficient wideband spectrum sensing," *Circuits and Systems I: Regular Papers, IEEE Trans.*, vol. 62, no. 2, pp. 527–535, 2015.

[15] R. T. Yazicigil, T. Haque, M. R. Whalen, J. Yuan, J. Wright, and P. R. Kinget, "Wideband rapid interferer detector exploiting compressed sampling with a quadrature analog-to-information converter," *Solid-State Circuits, IEEE Journal of*, vol. PP, no. 99, pp. 1–18, 2015.

[16] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, 1993.

[17] D. Ramasamy, S. Venkateswaran, and U. Madhow, "Compressive Parameter Estimation in AWGN," *Signal Processing, IEEE Trans.*, vol. 62, no. 8, pp. 2012–2027, 2014.

[18] P. Pakrooh, A. Pezeshki, L. L. Scharf, D. Cochran, and S. D. Howard, "Analysis of Fisher Information and the Cramer-Rao Bound for Nonlinear Parameter Estimation After Random Compression," *Signal Processing, IEEE Trans.*, vol. 63, no. 23, pp. 6423–6428, 2015.

[19] H. Cao, Y. T. Chan, and H. C. So, "Efficient sensing for compressive estimation of frequency of a real sinusoid," *Aerospace and Electronic Systems, IEEE Trans.*, pp. 1–1, 2020.

[20] P. Stoica and A. Nehorai, "Music, maximum likelihood and cramer-rao bound: further results and comparisons," in *International Conference on Acoustics, Speech, and Signal Processing,*, May 1989, pp. 2605–2608 vol.4.

[21] M. Marnat, M. Pelissier, O. Michel, and L. Ros, "Code Properties Analysis for the Implementation of a Modulated Wideband Converter," 2017, Proc. EUSIPCO.