ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/382601531

On the efficiency of misspecified Gaussian inference in nonlinear regression: Application to time-delay and Doppler estimation

Article *in* Signal Processing · July 2024

DOI: 10.1016/j.sigpro.2024.109614

CITATIONS
4

READS
18

2 authors:


Stefano Fortunati
Télécom SudParis (IP Paris) - SAMOVAR
82 PUBLICATIONS 1,267 CITATIONS

SEE PROFILE


Lorenzo Ortega Espluga
Institut Polytechnique des Sciences Avancées
63 PUBLICATIONS 319 CITATIONS

SEE PROFILE

All content following this page was uploaded by Lorenzo Ortega Espluga on 26 November 2024.

The user has requested enhancement of the downloaded file.

On the efficiency of misspecified Gaussian inference in nonlinear regression: application to time-delay and Doppler estimation

Stefano Fortunati[a,c,*], Lorenzo Ortega[b,c]

[a] *Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des signaux et systèmes, 91190, Gif-sur-Yvette*
[b] *TéSA, 7, Boulevard de la Gare, 31500, Toulouse, France*
[c] *DR2I - IPSA, France*

Abstract

Nonlinear regression plays a crucial role in various engineering applications. For the sake of mathematical tractability and ease of implementation, most of the existing inference procedures are derived under the assumption of *independent and identically distributed (i.i.d.) Gaussian-distributed* data. However, real-world situations often deviate from this assumption, with the true data generating process being a correlated, heavy-tailed and non-Gaussian one. The paper aims at providing the Misspecified Cramér-Rao Bound (MCRB) on the Mean Squared Error (MSE) of any unbiased (in a proper sense) estimator of the parameters of a nonlinear regression model derived under the i.i.d. Gaussian assumption in the place of the actual correlated, non-Gaussian data generating process. As a special case, the MCRB for an uncorrelated, i.i.d. Complex Elliptically Symmetric (CES) data generating process under Gaussian assumption is also provided. Consistency and asymptotic normality of the related Mismatched Maximum Likelihood Estimator (MMLE) will be discussed along with its connection with the Nonlinear Least Square Estimator (NLLSE) inherent to the nonlinear regression model. Finally, the derived theoretical findings will be applied in the well-known problem of time-delay and Doppler estimation for GNSS.

Keywords: Nonlinear regression, Misspecified Cramér-Rao bound, Mismatched Maximum Likelihood estimator, time-delay and Doppler estimation, band-limited signals.

1. Introduction

Nonlinear regression are one of the most-used statistical models in signal processing (SP) and related engineering applications. In a regression model, an observation vector

$$\mathbb{C}^N \ni \mathbf{x} = \mathbf{f}(\bar{\boldsymbol{\theta}}) + \mathbf{n}, \tag{1}$$

*Corresponding author
Email addresses: stefano.fortunati@centralesupelec.fr (Stefano Fortunati), lorenzo.ortega@tesa.prd.fr (Lorenzo Ortega)

is characterized by i) a vector of unknown deterministic parameters $\bar{\boldsymbol{\theta}}$, ii) a linear or nonlinear (continuous and differentiable) *known* function $\mathbf{f}$, parameterized by $\bar{\boldsymbol{\theta}}$ and iii) an additive noise vector $\mathbf{n}$. The function $\mathbf{f}$ generally specifies the measurement process while $\bar{\boldsymbol{\theta}}$ collects the quantities that need to be estimated. Regression models can be found in array processing, image processing, biomedical data analysis and even in climatic studies, just to name a few. While the definition of $\mathbf{f}$, $\bar{\boldsymbol{\theta}}$ and of the measurement noise $\mathbf{n}$ depends on the particular application at hands, the inference procedures used to estimate the parameter of interest usually share a common (although unrealistic) assumption: the entries of the noise vector $\mathbf{n}$ are sampled from an i.i.d. white Gaussian random process. This assumption is made to make the estimation algorithm mathematically tractable and easy to implement. In fact, it is well known that, under the i.i.d. Gaussian assumption, the optimal estimator is the nonlinear least square estimator (NLLSE). However, everyday engineering practice shows that this assumption is too simplistic since the noise process can be correlated and even non-Gaussian. The central question that we aim at answering in this paper is: *how accurate can an i.i.d., Gaussian-based inference procedure be when the regression model is characterized by a correlated, generally non-Gaussian noise?* In order to answer to this question, we will rely on the misspecification theory developed in [1, 2, 3, 4] and recently rediscovered in [5, 6, 7, 8, 9] and the references therein. In particular, in [4], a set of general conditions needed to guarantee the consistency and the asymptotic normality of the NLLSE for $\bar{\boldsymbol{\theta}}$ under depended data were provided along with the analytical expression of its error covariance matrix. Then, building upon the fundamental results presented in [4], in this paper we derive the so-called Misspecified Cramér-Rao Bound (MCRB) on the estimation of $\bar{\boldsymbol{\theta}}$ when the assumed model is the "classical" i.i.d. Gaussian model while the true data model is a dependent and non-Gaussian one. Moreover, we show that the error covariance matrix of the NLLSE, derived in [4], actually equates the proposed MCRB.

This paper is organized into seven distinct sections. In Sec. 2, we present both the true and assumed nonlinear regression models. Sec. 3 introduces the calculation of the pseudo-true parameter vector for the misspecified Gaussian nonlinear regression model. Sec. 4 derives the MCRB under quite general condition of the correlation structure of the true data generating process, while, in Sec. 5, we specialise this general results to a case of practical interest in which the true data model is an i.d.d. Complex Elliptical Symmetric (CES) model with unspecified density generator. Sec. 6 is dedicated to the investigation of the asymptotic properties of the NLLsE under the above mentioned misspecified scenario and to its relation with the MMLE. Sec. 7 provides an example of possible application of the theoretical results to the time-delay and Doppler estimation under the above-mentioned misspecified scenario for GNSS applications. Our conclusion are collected in Sec. 9. *Notation*: Throughout this paper, italics indicates scalar quantities (a), lower case and upper case boldface indicate column vectors ($\mathbf{a}$) and matrices ($\mathbf{A}$), respectively. Each entry of a matrix $\mathbf{A}$ is indicated as $a_{ij} \triangleq [\mathbf{A}]_{i,j}$. $\mathbf{I}_N$ defines the $N \times N$ identity matrix. The superscripts $*$, $\top$ and H indicate the complex conjugation, the transpose and the Hermitian operators respectively, then $\mathbf{A}^H = (\mathbf{A}^*)^\top$. The Euclidean norm of a vector $\mathbf{a}$ is indicated as $||\mathbf{a}||$.

2. Nonlinear regression with dependent observations

As discussed in the Introduction, the nonlinear regression is one of the most used statistical model in SP and statistics. The aim of this section is then to introduce firstly the model in its generality (i.e. *the true signal model*) and secondly to present its Gaussian-based, i.i.d. simplified version as it is generally assumed by SP practitioners for inference purposes.

2.1. True signal model

Let $\{x_k \in \mathbb{C}\}_{k=-\infty}^{+\infty}$ be a sequence of scalar, complex-valued, observations characterized by the following data generating process:

$$x_k = f_k(\bar{\boldsymbol{\theta}}) + n_k, \quad -\infty < k < +\infty, \tag{2}$$

where $\bar{\boldsymbol{\theta}} \in \Theta \subset \mathbb{R}^p$ indicates the real-valued [1], *true parameter vector* and Θ is a compact subset of $\mathbb{R}^p$. The functions $f_k : \Theta \to \mathbb{C}$, $-\infty < k < +\infty$ are *known continuous and differentiable* functions on Θ. In practical applications, the sequence (2) will be observed from a finite integer $N_1 \in \mathbb{Z}$ to a finite integer $N_2 \in \mathbb{Z}$, such that $-\infty < N_1 < N_2 < +\infty$. Consequently, by defining $N = |N_2 - N_1 + 1|$, the sequence (2) can be written in a vectorial form as in (1) where $\mathbf{n} \in \mathbb{C}^N$ is a *zero-mean* complex-valued noise vector whose N entries are assumed to be sampled from a Wide Sense Stationary (WSS) discrete random process $\{n_k : \forall k\}$ characterizing the measurement noise n_k in (2).

Let us now have a closer look at the statistical characterization of $\{n_k : \forall k\}$. As a *zero-mean* WSS process, $\{n_k : \forall k\}$ is fully characterised by (see e.g. [11, Sec. 15.5], [12, Sec. 1.3]): *i*) its autocorrelation function $r_n[k+j, k] = r_n[k+j-k] = r_n[j]$ and *ii*) the joint probability density function (pdf) of the N samples $\mathbb{C}^N \ni \mathbf{n} \sim p_{\mathbf{n}}$, for any values of N. For further reference, we indicate the marginal pdf of each sample as $n_k \sim p_{n_k}$. We make the following (non-Gaussian, non-i.i.d.) assumption:

Assumption 1. *Let $\{n_k : \forall k\}$ be a zero-mean, WSS discrete and circular complex-valued process [13] such that the joint pdf of N samples follows an unspecified pdf $\mathbf{n} \sim p_{\mathbf{n}}, \forall N$ admitting (at least) finite first and second order moments. Then, we assume that its autocorrelation function exists and satisfies $|r_n[j]| \triangleq |E_{p_{\mathbf{n}}}[n_{k+j}^* n_k]| = O(|j|^{-\gamma})$, $j \in \mathbb{Z}$, for some $\gamma > 0$ that controls the polynomial speed of decay to 0 of $|r_n[j]|$.*[2] *Note that the circularity of $\{n_k : \forall k\}$ implies that $E_{p_{\mathbf{n}}}[n_{k+j} n_k] = 0$, $\forall k, j$.*

The value of the positive scalar γ depends on the regularity conditions (in particular on the finiteness of the moments) that we are willing to accept on the process $\{n_k : \forall k\}$ and on the class C to which the joint

[1] We decided to work with real-valued parameters for two reasons. Firstly, in practical applications, the parameters of interest are real-valued physical quantities as e.g. time-delay and Doppler. Secondly, this choice allows us to avoid the technicalities related to the Wirtinger calculus [10] that may obscure the more important statistical concepts. It is worth stressing that this choice will not limit the generality of the derived results since any complex-valued vector can be recast in term of a real-valued vector by means of the standard isomorphism between $\mathbb{C}^p$ and $\mathbb{R}^{2p}$, i.e. $\mathbb{C}^p \ni \theta \leftrightarrows \tilde{\theta} \triangleq (\mathrm{Re}(\theta)^\top, \mathrm{Im}(\theta)^\top)^\top$

[2] Given a real-valued function $f(x)$ and a positive real-valued function $g(x)$, $f(x) = O(g(x))$ if and only if there exists a positive real number a and a real number x_0 such that $|f(x)| \leq a g(x)$, $\forall x \geq x_0$.

$p_{\mathbf{n}}$ belongs. For an in-depth discussion on this point we refer to [4, Sec. 2] an the references therein. In the rest of this paper, we limit ourselves to consider only random processes, with relevant joint pdf $p_{\mathbf{n}} \in \mathcal{C}$, that satisfy Assumption 1. It is worth stressing that, as detailed in [4, Sec. 2], this family of random processes is large enough to encompass the vast majority of (non-pathological) statistical models encountered in physics and engineering.

It is worth noticing here that, as a direct consequence of Assumption 1:

- the marginal pdf p_{n_k}, related to the joint $p_{\mathbf{n}} \in \mathcal{C}$, of each sample $n_k \sim p_{n_k}$ is left fully unspecified,
- the covariance matrix of $\mathbf{n}$, i.e. $\mathbf{\Sigma} \triangleq E_{p_{\mathbf{n}}}[\mathbf{n}\mathbf{n}^H]$, has the following Hermitian Toeplitz structure [12, Sec. 1.3]:

$$\begin{aligned}[\mathbf{\Sigma}]_{k,k+j} &= [\mathbf{\Sigma}]_{k+1,k+j+1} = [\mathbf{\Sigma}]^*_{k+j+1,k+1} \\ &= r_n[(k+j)-k] = r_n^*[k-(k+j)] = r_n[j] = r_n^*[-j],\end{aligned} \tag{3}$$

$$[\mathbf{\Sigma}]_{k,k} = r_n[0] = \bar{\sigma}_n^2, \forall k, \tag{4}$$

where $\bar{\sigma}_n^2$ is the true and generally unknown noise power.

We would like to stress that Assumption 1 allows for a wide range of realistic noise models [14]. In fact, we can note that any (Gaussian and non-Gaussian) stable *second-order stationary* (SOS) *AutoRegressive Moving Average* (ARMA) discrete process, of any finite orders (and with finite moments of sufficiently high order), satisfies Assumption 1, since the autocorrelation function of any stable SOS ARMA decays exponentially. It is well know that, by appropriately choosing the orders of the Autoregressive and of the Moving Average parts, an ARMA process can approximate the (continuous) power spectral density (PSD) of any complex discrete random processes [12, Ch. 3]. Moreover, a non-Gaussian ARMA can characterise the heavy-tailed behaviour of realistic noise models. Another popular noise model of practical interest satisfying Assumption 1 is the Compound-Gaussian (CG) (or *spherically invariant random vector* (SIRV)) model [15]. In fact, any SIRV $\mathbf{n} \in \mathbb{C}^N$ can be represented as [15, Def. 3] $\mathbf{n} =_d \sqrt{\tau}\mathbf{m}$ for some real-valued positive random variable τ, such that $E[\tau] = 1$, called *texture*, independent of the zero-mean, N-dimensional, circular, complex Gaussian random vector, called *speckle*, $\mathbf{m} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$, where $\mathbf{\Sigma}$ is the covariance matrix given in (3).

To conclude this section, we note that the pdf of the data vector in eq. (1) can be expressed as function of the unspecified noise pdf $p_{\mathbf{n}}$ as:

$$\mathbf{x} \sim p_{\bar{\boldsymbol{\epsilon}}} \triangleq p_{\bar{\boldsymbol{\epsilon}}}(\mathbf{x}; \bar{\sigma}_n^2, \bar{\boldsymbol{\theta}}) = p_{\mathbf{n}}(\mathbf{x} - \mathbf{f}(\bar{\boldsymbol{\theta}}); \bar{\sigma}_n^2), \tag{5}$$

where

$$\bar{\boldsymbol{\epsilon}} = (\bar{\sigma}_n^2, \bar{\boldsymbol{\theta}}^\top)^\top \in \Gamma \subset \mathbb{R}^+ \times \mathbb{R}^p \tag{6}$$

is the complete vector of the true parameters, where $\bar{\boldsymbol{\theta}}$ is the vector of the parameter of interest and $\mathbb{R}^+ \ni \bar{\sigma}_n^2 > 0$ a *nuisance parameter*, i.e. a term whose estimation is not strictly required but the lack of its knowledge may have an impact on the estimation performance of $\bar{\boldsymbol{\theta}}$.

2.2. Misspecified Gaussian, i.i.d. signal model

To do inference on the parameter vector $\bar{\boldsymbol{\theta}}$, and specifically to estimate it, a common procedures among SP practitioners is to assume a simplified model describing the statistical behaviour of the observations in the place of the true data generating process in (2). This *model misspecification* is dictated by two main reasons [5]. The first one is that the autocorrelation structure, as well as the pdf $p_{\mathbf{n}}$ is generally not a-priori known and not easy to obtain from physical considerations on the random experiment at hand. Secondly, one could prefer a simplified model in order to derive estimation algorithms that are easy to implement and fast to compute.

One of the most popular simplifying assumption is to consider the noise process $\{n_k : \forall k\}$ as a *zero-mean, White Gaussian WSS* random process. This implies that its autocorrelation function can be expressed as $r_n[j] = \bar{\sigma}_n^2 \delta[j]$, where $\delta[j]$ is the Kronecker delta sequence. As a consequence, the noise vector $\mathbf{n} \in \mathbb{C}^N$ is distributed as a centered complex normal random vector with diagonal covariance matrix, i.e. $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2 \boldsymbol{I}_N)$. This simplifying assumption leads to the following misspecified statistical model for the data vector $\mathbf{x} \in \mathbb{C}^N$ in (1):

$$\mathcal{F}_{\boldsymbol{\epsilon}} \triangleq \left\{ f_{\boldsymbol{\epsilon}} | f_{\boldsymbol{\epsilon}}(\mathbf{x}; \boldsymbol{\epsilon}) = \mathcal{CN}(\mathbf{f}(\boldsymbol{\theta}), \sigma_n^2 \boldsymbol{I}_N), \boldsymbol{\epsilon} \in \Gamma \right\}, \tag{7}$$

that is, each pdf belonging to $\mathcal{F}_{\boldsymbol{\epsilon}}$ can be expressed as:

$$f_{\boldsymbol{\epsilon}}(\mathbf{x}; \sigma_n^2, \boldsymbol{\theta}) = (\pi \sigma_n^2)^{-N} e^{-\frac{||\mathbf{x}-\mathbf{f}(\boldsymbol{\theta})||^2}{\sigma_n^2}}. \tag{8}$$

The question that we are going to answer in the next section is: *is it possible to derive a lower bound to the Mean Squared Error (MSE) of any unbiased, in the misspecified sense defined in [3, 6, 5] estimation procedure of $\bar{\boldsymbol{\theta}}$, derived under the Gaussian, i.i.d., misspecified model $\mathcal{F}_{\boldsymbol{\epsilon}}$ in* (7) *in the presence of dependent observations satisfying Assumption 1?*

To answer to this question, we evaluate the MCRB [5, 6, 7] on the estimation of $\bar{\boldsymbol{\theta}}$ when the assumed model is $\mathcal{F}_{\boldsymbol{\epsilon}}$ while the true data generating process is the (dependent, non-Gaussian) one introduced in (2). To this end, we start by evaluating the *pseudo-true parameter vector* $\boldsymbol{\epsilon}_0 \in \Gamma$, i.e. the vector in Γ that minimizes the Kullback-Leibler Divergence (KLD) [5, A1][6, Sec. 4.4.1] between the true (and unknown) pdf $\mathbf{x} \sim p_{\bar{\boldsymbol{\epsilon}}}$ and any element $f_{\boldsymbol{\epsilon}} \in \mathcal{F}_{\boldsymbol{\epsilon}}$ of the assumed misspecified model in (7). The vector $\boldsymbol{\epsilon}_0 \in \Gamma$ can be seen as a sort of "minimum divergence projector" of the true pdf onto the misspecified model $\mathcal{F}_{\boldsymbol{\epsilon}}$ and then it characterises the pdf $f_{\boldsymbol{\epsilon}_0} \in \mathcal{F}_{\boldsymbol{\epsilon}}$ closest to the true pdf $p_{\bar{\boldsymbol{\epsilon}}}$ in the KLD sense.

3. The pseudo-true parameter vector

As anticipated in the previous section, the pseudo-true parameter vector $\boldsymbol{\epsilon}_0$ is the element in the parameter space Γ that minimizes the KLD between the true data pdf $\mathbf{x} \sim p_{\bar{\boldsymbol{\epsilon}}}$ and any (possibly) misspecified pdf $f_{\boldsymbol{\epsilon}_0} \in \mathcal{F}_{\boldsymbol{\epsilon}}$ [3], [5, A1] and [6, Sec. 4.4.1]:

$$D(p_{\bar{\boldsymbol{\epsilon}}}||f_{\boldsymbol{\epsilon}}) = E_{p_{\bar{\boldsymbol{\epsilon}}}}\left[\ln\left(\frac{p_{\bar{\boldsymbol{\epsilon}}}(\mathbf{x};\bar{\boldsymbol{\epsilon}})}{f_{\boldsymbol{\epsilon}}(\mathbf{x};\boldsymbol{\epsilon})}\right)\right] \quad \mathbf{x} \sim p_{\bar{\boldsymbol{\epsilon}}},\ f_{\boldsymbol{\epsilon}} \in \mathcal{F}_{\boldsymbol{\epsilon}} \tag{9}$$

$E_{p_{\bar{\boldsymbol{\epsilon}}}}[\cdot]$ is the expectation with respect to (w.r.t.) the true model's pdf. Consequently:

$$\boldsymbol{\epsilon}_0 = \arg\min_{\boldsymbol{\epsilon}\in\Gamma} \{D(p_{\bar{\boldsymbol{\epsilon}}}||f_{\boldsymbol{\epsilon}})\} = \arg\min_{\boldsymbol{\epsilon}\in\Gamma} \left\{E_{p_{\bar{\boldsymbol{\epsilon}}}}[-\ln f_{\boldsymbol{\epsilon}}(\mathbf{x};\boldsymbol{\epsilon})]\right\}. \tag{10}$$

From (8), it follows directly that:

$$E_{p_{\bar{\boldsymbol{\epsilon}}}}[-\ln f_{\boldsymbol{\epsilon}}] = N\ln(\pi) + N\ln(\sigma_n^2) + \frac{E_{p_{\bar{\boldsymbol{\epsilon}}}}\left[||\mathbf{x}-\mathbf{f}(\boldsymbol{\theta})||^2\right]}{\sigma_n^2}. \tag{11}$$

By substituting (11) in (10), we have:

$$\begin{aligned}\boldsymbol{\epsilon}_0 &= \arg\min_{\boldsymbol{\epsilon}\in\Gamma} \left\{E_{p_{\bar{\boldsymbol{\epsilon}}}}\left[-\ln f_{\boldsymbol{\epsilon}}(\mathbf{x};\sigma_n^2,\boldsymbol{\theta})\right]\right\} \\ &= \arg\min_{\boldsymbol{\epsilon}\in\Gamma} \left\{E_{p_{\bar{\boldsymbol{\epsilon}}}}\left[\frac{1}{\sigma_n^2}\left[||\mathbf{x}-\mathbf{f}(\bar{\boldsymbol{\theta}})||^2\right]\right] + N\ln(\sigma_n^2)\right\}\end{aligned} \tag{12}$$

Since, as shown in the Appendix, the minimization w.r.t. to $\boldsymbol{\theta}$ is independent from the one of σ_n^2, $\boldsymbol{\theta}_0$ can be obtained as:

$$\boldsymbol{\theta}_0 = \arg\min_{\boldsymbol{\theta}} \left\{E_{p_{\bar{\boldsymbol{\epsilon}}}}[-\ln f_{\boldsymbol{\epsilon}}(\mathbf{x};\boldsymbol{\epsilon})]\right\} \tag{13a}$$

$$= \arg\min_{\boldsymbol{\theta}} \left\{E_{p_{\bar{\boldsymbol{\epsilon}}}}\left[\left[||\mathbf{x}-\mathbf{f}(\bar{\boldsymbol{\theta}})||^2\right]\right]\right\} \tag{13b}$$

$$= \arg\min_{\boldsymbol{\theta}} \left\{E_{p_{\bar{\boldsymbol{\epsilon}}}}\left[\mathrm{tr}\left((\mathbf{x}-\mathbf{f}(\bar{\boldsymbol{\theta}}))(\mathbf{x}-\mathbf{f}(\bar{\boldsymbol{\theta}}))^H\right)\right]\right\} \tag{13c}$$

$$= \arg\min_{\boldsymbol{\theta}} \left\{\mathrm{tr}\left(\boldsymbol{\Sigma} + \mathbf{f}(\bar{\boldsymbol{\theta}})\mathbf{f}(\bar{\boldsymbol{\theta}})^H - \mathbf{f}(\bar{\boldsymbol{\theta}})\mathbf{f}(\boldsymbol{\theta})^H - \mathbf{f}(\boldsymbol{\theta})\mathbf{f}(\bar{\boldsymbol{\theta}})^H + \mathbf{f}(\boldsymbol{\theta})\mathbf{f}(\boldsymbol{\theta})^H\right)\right\} \tag{13d}$$

$$= \arg\min_{\boldsymbol{\theta}} \left\{||\mathbf{f}(\boldsymbol{\theta}) - \mathbf{f}(\bar{\boldsymbol{\theta}})||^2\right\} \Rightarrow \boldsymbol{\theta}_0 = \bar{\boldsymbol{\theta}}. \tag{13e}$$

Remarkably, this result tells us that the pseudo-true parameter vector of interest $\boldsymbol{\theta}_0$ is equal to the one $\bar{\boldsymbol{\theta}}$.

Let us now minimize w.r.t. to the variance σ_n^2. By using the result obtained in (13), we have:

$$\sigma_0^2 = \arg\min_{\sigma_n^2} \left\{E_{p_{\bar{\boldsymbol{\epsilon}}}}\left[-\ln f_{\boldsymbol{\epsilon}}(\mathbf{x};\sigma_n^2,\bar{\boldsymbol{\theta}})\right]\right\} \tag{14}$$

$$\Rightarrow E_{p_{\bar{\epsilon}}}\left[\left.\frac{\partial}{\partial\sigma_n^2}\ln f_{\epsilon}(\mathbf{x};\sigma_n^2,\bar{\boldsymbol{\theta}})\right|_{\sigma_n^2=\sigma_0^2}\right]=0 \tag{15}$$

From direct calculation, we have:

$$E_{p_{\bar{\epsilon}}}\left[\left.\frac{\partial}{\partial\sigma_n^2}\ln f_{\epsilon}(\mathbf{x};\sigma_n^2,\bar{\boldsymbol{\theta}})\right|_{\sigma_n^2=\sigma_0^2}\right] \tag{16a}$$

$$= E_{p_{\bar{\epsilon}}}\left[\left.-\frac{N}{\sigma_n^2}+\frac{1}{\sigma_n^4}\|\mathbf{x}-\mathbf{f}(\bar{\boldsymbol{\theta}})\|^2\right|_{\sigma_n^2=\sigma_0^2}\right] \tag{16b}$$

$$= E_{p_{\bar{\epsilon}}}\left[-\frac{N}{\sigma_n^2}+\frac{\mathrm{tr}(\mathbf{n}\mathbf{n}^H)}{\sigma_n^4}\right]=-\frac{N}{\sigma_0^2}+\frac{\mathrm{tr}\,(\boldsymbol{\Sigma})}{\sigma_0^4} \tag{16c}$$

$$= -\frac{N}{\sigma_0^2}+\frac{Nr_n[0]}{\sigma_0^4}=-\frac{N}{\sigma_0^2}+\frac{N\bar{\sigma}_n^2}{\sigma_0^4}=0\Rightarrow\sigma_0^2=\bar{\sigma}_n^2 \tag{16d}$$

Again, eq. (16) tells us that the pseudo-true nuisance parameter σ_0^2 equates the true one $\bar{\sigma}_n^2$.

By collecting the results from eqs. (13) and (16), we have that the pseudo-true parameter vector equates the true one

$$\boldsymbol{\epsilon}_0=\bar{\boldsymbol{\epsilon}}\triangleq(\bar{\sigma}_n^2,\bar{\boldsymbol{\theta}}^\top)^\top, \tag{17}$$

under mild assumptions, i.e. for any noise vector $\mathbb{C}^N\ni\mathbf{n}\sim p_{\mathbf{n}}$ sampled form a discrete random process $\{n_k:\forall k\}$ whose unspecified joint pdf has finite first and second order moments, that is it admits a zero-mean $E_{p_{\mathbf{n}}}[\mathbf{n}]=\mathbf{0}$ and a covariance matrix $\boldsymbol{\Sigma}\triangleq E_{p_{\mathbf{n}}}[\mathbf{n}\mathbf{n}^H]$ satisfying (3) and (4). It can be noted that the equality in (17) does not requires the polynomial decrease of the autocorrelation function introduced in Assumption 1. However, we will see that this requirement will be crucial to derive asymptotic results about the efficiency of misspecifed Gaussian procedures.

4. Closed form expression for the MCRB

The aim of this section is to provide the closed form expression of the Misspecified Cramér-Rao Bound (MCRB) for the estimation of $\bar{\boldsymbol{\epsilon}}$ under the misspecified scenario discussed in Sec. 2.2. Following [3], [5, Theo. 1] and [6, Theo. 4.1] and by exploiting the equality between the true and the pseudo-true paramater vectors, the MCRB is given by:

$$\mathbf{MCRB}(\boldsymbol{\epsilon}_0)=\mathbf{MCRB}(\bar{\boldsymbol{\epsilon}})=\mathbf{A}(\bar{\boldsymbol{\epsilon}})^{-1}\mathbf{B}(\bar{\boldsymbol{\epsilon}})\mathbf{A}(\bar{\boldsymbol{\epsilon}})^{-1}, \tag{18}$$

where:

$$[\mathbf{A}(\bar{\boldsymbol{\epsilon}})]_{i,j}\triangleq\left[E_{p_{\bar{\epsilon}}}\left[\nabla_{\epsilon}\nabla_{\epsilon}^\top\ln f_{\epsilon}(\mathbf{x};\bar{\boldsymbol{\epsilon}})\right]\right]_{i,j}=E_{p_{\bar{\epsilon}}}\left[\left.\frac{\partial^2}{\partial_i\partial_j}\ln f_{\epsilon}(\mathbf{x};\boldsymbol{\epsilon})\right|_{\boldsymbol{\epsilon}=\bar{\boldsymbol{\epsilon}}}\right], \tag{19}$$

$$[\mathbf{B}(\bar{\boldsymbol{\epsilon}})]_{i,j} \triangleq \left[E_{p_{\bar{\epsilon}}}\left[\nabla_{\boldsymbol{\epsilon}} \ln f_{\boldsymbol{\epsilon}}(\mathbf{x};\bar{\boldsymbol{\epsilon}})\nabla_{\boldsymbol{\epsilon}}^{\top} \ln f_{\boldsymbol{\epsilon}}(\mathbf{x};\bar{\boldsymbol{\epsilon}})\right]\right]_{i,j} = E_{p_{\bar{\epsilon}}}\left[\frac{\partial}{\partial_i} \ln f_{\boldsymbol{\epsilon}}(\mathbf{x};\boldsymbol{\epsilon})\bigg|_{\boldsymbol{\epsilon}=\bar{\boldsymbol{\epsilon}}} \frac{\partial}{\partial_j} \ln f_{\boldsymbol{\epsilon}}(\mathbf{x};\boldsymbol{\epsilon})\bigg|_{\boldsymbol{\epsilon}=\bar{\boldsymbol{\epsilon}}}\right], \tag{20}$$

where $f_{\boldsymbol{\epsilon}}(\mathbf{x};\boldsymbol{\epsilon}) \in \mathcal{F}_{\boldsymbol{\epsilon}}$ in (7).

The calculation of the matrices $\mathbf{A}(\bar{\boldsymbol{\epsilon}})$ and $\mathbf{B}(\bar{\boldsymbol{\epsilon}})$ will be performed in four steps:

1. *Evaluation of the terms related to* $\bar{\sigma}_n^2$.

   Through direct calculation, we have:

$$\begin{aligned} \nabla_{\sigma_n^2} \ln f_{\boldsymbol{\epsilon}}(\mathbf{x};\bar{\boldsymbol{\epsilon}}) &= \frac{\partial \ln f_{\boldsymbol{\epsilon}}(\mathbf{x};\bar{\sigma}_n^2,\bar{\boldsymbol{\theta}})}{\partial \sigma_n^2} \\ &= -\frac{N}{\sigma_n^2} + \frac{1}{\sigma_n^4}\|\mathbf{x}-\mathbf{f}(\bar{\boldsymbol{\theta}})\|^2\bigg|_{\sigma_n^2=\bar{\sigma}_n^2} = -\frac{N}{\bar{\sigma}_n^2} + \frac{\mathrm{tr}(\mathbf{n}\mathbf{n}^H)}{\bar{\sigma}_n^4}, \end{aligned} \tag{21}$$

   and then:

$$\begin{aligned} \nabla_{\sigma_n^2}\nabla_{\sigma_n^2}^{\top} \ln f_{\boldsymbol{\epsilon}}(\mathbf{x};\bar{\boldsymbol{\epsilon}}) &= \frac{N}{\sigma_n^4} - \frac{2}{\sigma_n^6}\|\mathbf{x}-\mathbf{f}(\bar{\boldsymbol{\theta}})\|^2\bigg|_{\sigma_n^2=\bar{\sigma}_n^2} \\ &= \frac{N}{\bar{\sigma}_n^4} - \frac{2\mathrm{tr}(\mathbf{n}\mathbf{n}^H)}{\bar{\sigma}_n^6}. \end{aligned} \tag{22}$$

   By taking the expectation w.r.t. the true data distribution $p_{\bar{\epsilon}}$ and following the same calculation done in (16), we get:

$$E_{p_{\bar{\epsilon}}}\left[\nabla_{\sigma_n^2}\nabla_{\sigma_n^2}^{\top} \ln f_{\boldsymbol{\epsilon}}(\mathbf{x};\bar{\boldsymbol{\epsilon}})\right] = \frac{N}{\bar{\sigma}_n^4} - \frac{2\mathrm{tr}\,(\boldsymbol{\Sigma})}{\bar{\sigma}_n^6} = -\frac{N}{\bar{\sigma}_n^4}, \tag{23}$$

   where we used their linearity to invert the order of the expectation and trace operators. Similarly, we have that;

$$\begin{aligned} E_{p_{\bar{\epsilon}}}\left[\left(\nabla_{\sigma_n^2} \ln f_{\boldsymbol{\epsilon}}(\mathbf{x};\bar{\boldsymbol{\epsilon}})\right)^2\right] &= E_{p_{\mathbf{n}}}\left[\left(-\frac{N}{\bar{\sigma}_n^2} + \frac{\mathbf{n}^H\mathbf{n}}{\bar{\sigma}_n^4}\right)^2\right] \\ &= \frac{N^2}{\bar{\sigma}_n^4} - \frac{2N\mathrm{tr}\,(\boldsymbol{\Sigma})}{\bar{\sigma}_n^6} + \frac{E_{p_{\mathbf{n}}}\left[(\mathbf{n}^H\mathbf{n})^2\right]}{\bar{\sigma}_n^8} \\ &= \frac{E_{p_{\mathbf{n}}}\left[(\mathbf{n}^H\mathbf{n})^2\right] - \bar{\sigma}_n^4 N^2}{\bar{\sigma}_n^8}. \end{aligned} \tag{24}$$

   Note that the term $E_{p_{\mathbf{n}}}\left[(\mathbf{n}^H\mathbf{n})^2\right]$ cannot be further developed without specifying the true pdf of the noise $p_{\mathbf{n}}$. We will further discuss this point in the next section.

2. *Evaluation of the terms related to* $\bar{\boldsymbol{\theta}}$

   From the assumed Gaussian pdf in eq. (8), we have:

$$\nabla_{\boldsymbol{\theta}} \ln f_{\boldsymbol{\epsilon}}(\mathbf{x};\bar{\sigma}_n^2,\bar{\boldsymbol{\theta}}) = -\frac{1}{\bar{\sigma}_n^2}\nabla_{\boldsymbol{\theta}}\|\mathbf{x}-\mathbf{f}(\bar{\boldsymbol{\theta}})\|^2$$

$$= \frac{1}{\bar{\sigma}_n^2} \sum_{k=N_1}^{N_2} \left[\left(x_k - f_k(\bar{\boldsymbol{\theta}})\right)^* \nabla_{\boldsymbol{\theta}} \bar{f}_k + \left(x_k - f_k(\bar{\boldsymbol{\theta}})\right) \nabla_{\boldsymbol{\theta}}^* \bar{f}_k \right]$$

$$= \frac{2}{\bar{\sigma}_n^2} \sum_{k=N_1}^{N_2} \mathrm{Re}\left\{ n_k^* \nabla_{\boldsymbol{\theta}} \bar{f}_k \right\}, \quad (25)$$

where, for ease of notation, we posed $\nabla_{\boldsymbol{\theta}} f_k(\bar{\boldsymbol{\theta}}) = \nabla_{\boldsymbol{\theta}} \bar{f}_k$.

According to the eq. (20), we can evaluate the matrix $\mathbf{B}(\bar{\boldsymbol{\epsilon}})$ as showed in eq. (26) reported at the bottom of this page. It is worth noticing that, in the step (26d), we used the circularity assumption on $\{n_k : \forall k\}$, i.e. $E_{p_n}[n_k n_j] = 0$, $\forall k, j$ (see Assumption 1). The matrix $\mathbf{P}(\bar{\boldsymbol{\theta}})$ in (26g) has been introduced for further reference.

Moreover, again through direct calculation, we have :

$$\nabla_{\boldsymbol{\theta}} \nabla_{\boldsymbol{\theta}}^\top \ln f_{\boldsymbol{\epsilon}}(\mathbf{x}; \bar{\sigma}_n^2, \bar{\boldsymbol{\theta}})$$

$$= \frac{1}{\bar{\sigma}_n^2} \sum_{k=N_1}^{N_2} \left[\left(x_k - f_k(\bar{\boldsymbol{\theta}})\right)^* \nabla_{\boldsymbol{\theta}} \nabla_{\boldsymbol{\theta}}^\top \bar{f}_k - \nabla_{\boldsymbol{\theta}} \bar{f}_k \nabla_{\boldsymbol{\theta}}^H \bar{f}_k \right] + \frac{1}{\bar{\sigma}_n^2} \sum_{k=N_1}^{N_2} \left[\left(x_k - f_k(\bar{\boldsymbol{\theta}})\right) \left[\nabla_{\boldsymbol{\theta}} \nabla_{\boldsymbol{\theta}}^\top \bar{f}_k\right]^* - \nabla_{\boldsymbol{\theta}}^* \bar{f}_k \nabla_{\boldsymbol{\theta}}^\top \bar{f}_k \right]$$

$$= \frac{2}{\bar{\sigma}_n^2} \sum_{k=N_1}^{N_2} \mathrm{Re}\left\{ \left(x_k - f_k(\bar{\boldsymbol{\theta}})\right) \left[\nabla_{\boldsymbol{\theta}} \nabla_{\boldsymbol{\theta}}^\top \bar{f}_k\right]^* \right\} - \frac{2}{\bar{\sigma}_n^2} \sum_{k=N_1}^{N_2} \mathrm{Re}\left\{ \nabla_{\boldsymbol{\theta}} \bar{f}_k \nabla_{\boldsymbol{\theta}}^H \bar{f}_k \right\}. \quad (27a)$$

Then we can introduce the matrix $\mathbf{A}(\bar{\boldsymbol{\theta}})$ as

$$\mathbf{A}(\bar{\boldsymbol{\theta}}) \triangleq E_{p_{\boldsymbol{\epsilon}}}\left[\nabla_{\boldsymbol{\theta}} \nabla_{\boldsymbol{\theta}}^\top \ln f_{\boldsymbol{\epsilon}}(\mathbf{x}; \bar{\sigma}_n^2, \bar{\boldsymbol{\theta}}) \right] \quad (28)$$

$$= -\frac{N}{\bar{\sigma}_n^2} \left[\frac{2}{N} \sum_{k=N_1}^{N_2} \mathrm{Re}\left\{ \nabla_{\boldsymbol{\theta}} \bar{f}_k \nabla_{\boldsymbol{\theta}}^H \bar{f}_k \right\} \right] \triangleq -\frac{N}{\bar{\sigma}_n^2} \mathbf{K}(\bar{\boldsymbol{\theta}}),$$

where, again, we have introduced the matrix $\mathbf{K}(\bar{\boldsymbol{\theta}})$ for further reference. Note that the expectation of the term in (27a) is nil since $E_{p_{\bar{\boldsymbol{\epsilon}}}}\left[x_k - f_k(\bar{\boldsymbol{\theta}})\right] = E_{p_{\bar{\boldsymbol{\epsilon}}}}[n_k] = 0, \forall k$.

3. *Evaluation of the cross-terms*

From the circularity of the noise process (see Assumption 1), it is immediate to verify that:

$$E_{p_{\boldsymbol{\epsilon}}}\left[\nabla_{\sigma_n^2} \ln f_{\boldsymbol{\epsilon}}(\mathbf{x}; \bar{\boldsymbol{\epsilon}}) \nabla_{\boldsymbol{\theta}}^\top \ln f_{\boldsymbol{\epsilon}}(\mathbf{x}; \bar{\sigma}_n^2, \bar{\boldsymbol{\theta}}) \right] =$$

$$E_{p_{\boldsymbol{\epsilon}}}\left[\nabla_{\sigma_n^2} \ln f_{\boldsymbol{\epsilon}}(\mathbf{x}; \bar{\boldsymbol{\epsilon}}) \nabla_{\boldsymbol{\theta}} \ln f_{\boldsymbol{\epsilon}}(\mathbf{x}; \bar{\sigma}_n^2, \bar{\boldsymbol{\theta}}) \right]^\top = \mathbf{0}_{1 \times p}. \quad (29)$$

Moreover, we have that:

$$\nabla_{\sigma_n^2} \nabla_{\boldsymbol{\theta}} \ln f_{\boldsymbol{\epsilon}}(\mathbf{x}; \bar{\sigma}_n^2, \bar{\boldsymbol{\theta}}) = \nabla_{\boldsymbol{\theta}} \nabla_{\sigma_n^2} \ln f_{\boldsymbol{\epsilon}}(\mathbf{x}; \bar{\sigma}_n^2, \bar{\boldsymbol{\theta}})$$

$$= \frac{2}{\bar{\sigma}_n^4} \sum_{k=N_1}^{N_2} \mathrm{Re}\left\{ n_k^* \nabla_{\boldsymbol{\theta}} \bar{f}_k \right\}. \quad (30)$$

Consequently, since the noise process is zero-mean, we trivially have that:

$$
\begin{aligned}
&E_{p_{\bar{\epsilon}}}\left[\nabla_{\sigma_n^2}\nabla_{\boldsymbol{\theta}}^\top \ln f_{\boldsymbol{\epsilon}}(\mathbf{x};\bar{\sigma}_n^2,\bar{\boldsymbol{\theta}})\right] \\
&\quad = E_{p_{\boldsymbol{\epsilon}}}\left[\nabla_{\boldsymbol{\theta}}\nabla_{\sigma_n^2} \ln f_{\boldsymbol{\epsilon}}(\mathbf{x};\bar{\sigma}_n^2,\bar{\boldsymbol{\theta}})\right]^\top = \mathbf{0}_{1\times p}.
\end{aligned} \tag{31}
$$

4. *Definition of the matrices* $\mathbf{A}(\bar{\boldsymbol{\epsilon}})$ *and* $\mathbf{B}(\bar{\boldsymbol{\epsilon}})$

By collecting the previous results, we have that the matrix $\mathbf{A}(\bar{\boldsymbol{\epsilon}})$ in eq. (19) can be expressed as:

$$
\mathbf{A}(\bar{\boldsymbol{\epsilon}}) = N\begin{pmatrix} -1/\bar{\sigma}_n^4 & \mathbf{0}_{1\times p} \\ \mathbf{0}_{p\times 1} & -\frac{1}{\bar{\sigma}_n^2}\mathbf{K}(\bar{\boldsymbol{\theta}}) \end{pmatrix}. \tag{32}
$$

Similarly, for the matrix $\mathbf{B}(\bar{\boldsymbol{\epsilon}})$ in eq. (20), we have:

$$
\mathbf{B}(\bar{\boldsymbol{\epsilon}}) = N\begin{pmatrix} \frac{(E_{p_{\mathbf{n}}}\left[(\mathbf{n}^H\mathbf{n})^2\right]-\bar{\sigma}_n^4N^2)}{N\bar{\sigma}_n^8} & \mathbf{0}_{1\times p} \\ \mathbf{0}_{p\times 1} & \frac{1}{\bar{\sigma}_n^4}\mathbf{P}(\bar{\boldsymbol{\theta}}) \end{pmatrix}. \tag{33}
$$

As we can see from eq. (33), the matrix $\mathbf{B}(\bar{\boldsymbol{\epsilon}})$ is function of the matrix $\mathbf{P}(\bar{\boldsymbol{\theta}})$ in (26g) and the term $E_{p_{\mathbf{n}}}[(\mathbf{n}^H\mathbf{n})^2]$. Now, in order to provide asymptotic results on the number of observations, i.e. as $N\to\infty$, the norm of the matrix matrix $\mathbf{B}(\bar{\boldsymbol{\epsilon}})$ has to remain bounded as $N\to\infty$. As discussed in [4], the polynomial decrease of $r_n[j]$ is needed to guarantees that $\mathbf{B}(\bar{\boldsymbol{\epsilon}})$ will not explode as $N\to\infty$. [3] Finally, the MCRB in (18) can be expressed as:

$$
\begin{aligned}
\mathbf{MCRB}(\bar{\boldsymbol{\epsilon}}) &= \mathbf{A}(\bar{\boldsymbol{\epsilon}})^{-1}\mathbf{B}(\bar{\boldsymbol{\epsilon}})\mathbf{A}(\bar{\boldsymbol{\epsilon}})^{-1} \\
&= \frac{1}{N}\begin{pmatrix} (E_{p_{\mathbf{n}}}\left[(\mathbf{n}^H\mathbf{n})^2\right]-\bar{\sigma}_n^4N^2)/N & \mathbf{0}_{1\times p} \\ \mathbf{0}_{p\times 1} & \mathbf{C}(\bar{\boldsymbol{\theta}}) \end{pmatrix},
\end{aligned} \tag{34}
$$

where

$$
\mathbf{C}(\bar{\boldsymbol{\theta}}) \triangleq \mathbf{K}(\bar{\boldsymbol{\theta}})^{-1}\mathbf{P}(\bar{\boldsymbol{\theta}})\mathbf{K}(\bar{\boldsymbol{\theta}})^{-1}. \tag{35}
$$

It is important to note that, due to the block-diagonal structure of $\mathbf{MCRB}(\bar{\boldsymbol{\epsilon}})$, the MCRB of the parameter of interest vector $\bar{\boldsymbol{\theta}}$ can be simply obtained as:

$$
\mathbf{MCRB}(\bar{\boldsymbol{\theta}}) = N^{-1}\mathbf{C}(\bar{\boldsymbol{\theta}}). \tag{36}
$$

Remarkably, this result tells us that the estimation of $\bar{\boldsymbol{\theta}}$ is asymptotically decorrelatd from the nuisance parameter $\bar{\sigma}_n^2$.

[3] Roughly speaking, the Assumption 1 guarantees the existence of a matrix $\mathbf{B}^0$, such that $\det(\mathbf{B}^0) > 0$ and $\mathbf{a}^\top(\mathbf{B}^0 - \mathbf{B}(\bar{\boldsymbol{\epsilon}}))\mathbf{a}\to 0$ as $N\to\infty$, for any non-zero real vector $\mathbf{a}\in\mathbb{R}^{p+1}$.

4.1. Consistent estimation of the matrix $\mathbf{P}(\bar{\boldsymbol{\theta}})$

Let us take a closer look to the matrix $\mathbf{P}(\bar{\boldsymbol{\theta}})$ in eq. (26f). It can be immediately noted that it depends on the a-priori knowledge of the autocorrelation function of the noise $r_n[j] \triangleq E_{p_{\mathbf{n}}}[n^*_{k+j} n_k]$. However, to evaluate it, we need to know the true pdf $p_{\mathbf{n}}$ of the noise. This is in contrast with the Assumption 1 where $p_{\mathbf{n}} \in C$ is left unspecified. We should then rely on a consistent estimator $\widehat{\mathbf{P}}_N$ of $\mathbf{P}(\bar{\boldsymbol{\theta}})$. Thanks to Assumption 1, deriving such consistent estimator is possible, even in presence of dependent observations. Following [4], let us define the estimator $\widehat{\mathbf{P}}_N$ of $\mathbf{P}(\bar{\boldsymbol{\theta}})$ as:

$$\widehat{\mathbf{P}}_N = \frac{2}{N} \sum\nolimits_{k=N_1}^{N_2} |\hat{n}_k|^2 \mathrm{Re}\left\{ \nabla_{\boldsymbol{\theta}} \hat{f}_k \nabla^H_{\boldsymbol{\theta}} \hat{f}_k \right\} + \frac{4}{N} \sum\nolimits_{j=1}^{l} \sum\nolimits_{k=N_1}^{N_2-j} \mathrm{Re}\left\{ \hat{n}^*_{k+j} \hat{n}_k \nabla_{\boldsymbol{\theta}} \hat{f}_{k+j} \nabla^H_{\boldsymbol{\theta}} \hat{f}_k \right\}, \tag{37}$$

where $l \leqslant N_2 - N_1 + 1$ is the so-called *correlation lag* formally defined in Theo. 1, $\hat{n}_k \triangleq x_k - f_k(\hat{\boldsymbol{\theta}}_N)$ and $\nabla_{\boldsymbol{\theta}} \hat{f}_k \triangleq \nabla_{\boldsymbol{\theta}} f_k(\hat{\boldsymbol{\theta}}_N)$ and $\hat{\boldsymbol{\theta}}_N$ is a $\sqrt{N}$-consistent estimator of the true parameter vector $\bar{\boldsymbol{\theta}}$. Among all the possible consistent estimators, the best choice is the asymptotic efficient one that we are going to introduce in the subsequent Sec. 6. The consistency of the estimator $\widehat{\mathbf{P}}_N$ in (37) is established in [4, Theo. 3.5]:

Theorem 1. *Under Assumption 1 and other technical regularity conditions (see A1a, A3, A4 and A7 in [4]), if the correlation lag* l *grows at the rate* $l = o(N^{1/3})$ *as* $N \to \infty$,[4] *we have that* $\hat{\mathbf{P}}_N$ *is a consistent estimator of* $\mathbf{P}(\bar{\boldsymbol{\theta}})$*:*

$$\widehat{\mathbf{P}}_N \xrightarrow{p} \mathbf{P}_N(\bar{\boldsymbol{\theta}}), \tag{38}$$

where $\xrightarrow{p}$ *indicates the convergence (element by element) in probability.*

By a direct application of the Continuous Mapping Theorem and of the Slutsky's Lemma [16, Theo. 2.3 and Lemma 2.8], we have that the matrix $\mathbf{C}(\bar{\boldsymbol{\theta}})$ in (35) can be consistently estimated as:

$$\widehat{\mathbf{C}}_N \triangleq \mathbf{K}(\hat{\boldsymbol{\theta}}_N)^{-1} \widehat{\mathbf{P}}_N \mathbf{K}(\hat{\boldsymbol{\theta}}_N)^{-1} \xrightarrow{p} \mathbf{C}(\bar{\boldsymbol{\theta}}), \tag{39}$$

that can be exploited to get a consistent estimation of the MCRB on the vector of the parameters of interest $\bar{\boldsymbol{\theta}}$.

5. MCRB expression for CES uncorrelated data

In order to highlight the importance and the generality of the results obtained in the previous section, let us consider the particular case, yet important in applications, where the noise process $\{n_k : \forall k\}$ is assumed to be *white* with *independent* and *identically Complex Elliptically Symmetric (CES)-distributed* samples. More formally, we assume that:

[4] Given a real-valued function $f(x)$ and a strictly positive real-valued function $g(x)$, $f(x) = o(g(x))$ if for every positive real number a, there exists a real number x_0 such that $|f(x)| \le a g(x)$, $\forall x \ge x_0$.

Assumption 2. *Let $\{n_k : \forall k\}$ be a zero-mean, white WSS discrete and circular complex-valued process [13] such that:*

1. *each sample n_k follows a CES distribution $n_k \sim p_n = CES(0, \bar{\sigma}_n^2, g)$ with unspecified density generator g,*
2. *its autocorrelation satisfies $r_n[j] \triangleq E_{p_n}[n_{k+j}^* n_k] = \bar{\sigma}_n^2 \delta[j]$, where $\delta[j]$ is the Kronecker delta sequence.*

As a direct consequence of Assumption 2, we have that:

- the joint pdf $p_{\mathbf{n}}$ of the noise vector $\mathbf{n} \sim p_{\mathbf{n}}$ is the product of the marginal densities, i.e. $p_{\mathbf{n}}(\mathbf{n}; \bar{\sigma}_n^2, g) = \prod_{k=N_1}^{N_2} p_{n_k}(n_k; \bar{\sigma}_n^2, g)$
- the covariance matrix of $\mathbf{n} \in \mathbb{C}^N$ is a diagonal matrix, i.e. $\mathbf{\Sigma} \triangleq E_{p_n}[\mathbf{n}\mathbf{n}^H] = \bar{\sigma}_n^2 \mathbf{I}_N$.

It is worth stressing here the generality of an *unspecified CES distribution* for the noise samples. The CES ones is a wide class of non-Gaussian and heavy-tailed distributions encompassing the Gaussian, the Generalized Gaussian, the $t-$, the $K-$ and the Weibull distributions as special cases [15]. Since its nominal density generator g is left unspecified, we let the noise $n_k \sim p_{n_k}$ have any possible distribution in the CES class.

From the Stochastic Representation Theorem [15, Theo. 3], each entry n_k can be represented as [15, Theo. 3]:

$$n_k =_d \sqrt{Q_k} \bar{\sigma}_n u_k, \tag{40}$$

where u_k is a complex univariate random variable uniformly distributed on $\mathbb{C}S \triangleq \{u \in \mathbb{C} | |u| = 1\}$, i.e. $u_k \sim U(\mathbb{C}S)$. The *second order modular variate* $Q_k \sim Q$ is a positive random variable, independent from u_k with pdf $p_Q(q) = \delta_g^{-1} g(q)$, where $\delta_g \triangleq \int_0^\infty g(q)dq$ is a normalizing constant (see [15, Eq. (19)]). Since the density generator g is left unspecified, it is immediate to verify that there is a scale ambiguity between $\bar{\sigma}_n^2$ and g itself. To avoid this problem, we impose that $E[Q] = 1$. Note that, this constraint allows us to consider $\bar{\sigma}_n^2$ as the *statistical power* P of the noise n_k, (see the discussion in [15, Sec. III.C]), since from (40), we have that:

$$P \triangleq E[|n_k|^2] = E[Q]E[|u_k|^2]\bar{\sigma}_n^2 = \bar{\sigma}_n^2, \tag{41}$$

where $E[|u_k|^2] = 1$ [15, Lemma 1].

Let us now apply the general expression of the MCRB obtained in (34) to the special case of an i.i.d. CES-distributed noise process formally characterized in Assumption 2. To this end, it is immediate to verify that the matrix $\mathbf{A}(\bar{\boldsymbol{\epsilon}})$ in (32) remain unchanged. Let us now focus on the matrix $\mathbf{B}(\bar{\boldsymbol{\epsilon}})$ in (33). We will proceed as follows.

1. *Evaluation of the term* $[\mathbf{B}(\bar{\boldsymbol{\epsilon}})]_{1,1}$ *in* (24).

Let us start by evaluating the term $E_{p_\mathbf{n}}\left[(\mathbf{n}^H\mathbf{n})^2\right]$. Under Assumption 2 and by exploiting the stochastic representation in (40), the term $E_{p_\mathbf{n}}[(\mathbf{n}^H\mathbf{n})^2]$ can be evaluated as shown in (42), reported at the bottom of the page,where we used:

- the mutual independence between $Q_i \sim Q$ and $Q_j \sim Q$ and between u_i and u_j (see Assumption 2),
- from the i.i.d. assumption, we have that $E\left[Q_i^2\right] = E\left[Q^2\right], \forall i$,
- the constraint $E\left[Q_i\right] = E\left[Q_j\right] = E\left[Q\right] = 1$,
- the relations $E\left[|u_j|^2\right] = 1$ and $E\left[|u_j|^4\right] = 1$ from [15, Lemma 1].

By using this result, the term $[\mathbf{B}(\bar{\boldsymbol{\epsilon}})]_{1,1}$ can be readily expressed as:

$$[\mathbf{B}_{iid}(\bar{\boldsymbol{\epsilon}})]_{1,1} = N(E\left[Q^2\right] - 1)/\bar{\sigma}_n^4 \tag{43}$$

2. *Evaluation of the matrix* $\mathbf{B}(\bar{\boldsymbol{\epsilon}})$

By putting in the general expression of $\mathbf{B}(\bar{\boldsymbol{\theta}})$ in (26) the autocorrelation function $r_n[j] = \bar{\sigma}_n^2\delta[j]$ (see Assumption 2), it is immediate to verify that:

$$\mathbf{B}_{iid}(\bar{\boldsymbol{\theta}}) = \frac{2}{\bar{\sigma}_n^2}\sum_{k=N_1}^{N_2}\mathrm{Re}\left\{\nabla_{\boldsymbol{\theta}}\bar{f}_k\nabla_{\boldsymbol{\theta}}^H\bar{f}_k\right\} \triangleq \frac{N}{\bar{\sigma}_n^2}\mathbf{K}(\bar{\boldsymbol{\theta}}), \tag{44}$$

where $\mathbf{K}(\bar{\boldsymbol{\theta}})$ is the matrix already defined eq. (28).

Consequently, the matrix $\mathbf{B}(\bar{\boldsymbol{\epsilon}})$, under Assuption 2, can be expressed as:

$$\mathbf{B}_{iid}(\bar{\boldsymbol{\epsilon}}) = N\begin{pmatrix} (E[Q^2]-1)/\sigma_n^4 & \mathbf{0}_{1\times 4} \\ \mathbf{0}_{4\times 1} & \frac{1}{\bar{\sigma}_n^2}\mathbf{K}(\bar{\boldsymbol{\theta}}) \end{pmatrix}. \tag{45}$$

Finally, from the general expression in (34), the MCRB for the estimation of $\bar{\boldsymbol{\epsilon}}$ under Assumption 2 can be expressed as:

$$\mathbf{MCRB}_{iid}(\bar{\boldsymbol{\epsilon}}) = \mathbf{A}(\bar{\boldsymbol{\epsilon}})^{-1}\mathbf{B}_{iid}(\bar{\boldsymbol{\epsilon}})\mathbf{A}(\bar{\boldsymbol{\epsilon}})^{-1} = \frac{1}{N}\begin{pmatrix} \sigma_n^4(E[Q^2]-1) & \mathbf{0}_{1\times 4} \\ \mathbf{0}_{4\times 1} & \bar{\sigma}_n^2\mathbf{K}(\bar{\boldsymbol{\theta}})^{-1} \end{pmatrix}, \tag{46}$$

and consequently, due to the block-diagonal structure of $\mathbf{MCRB}_{iid}(\bar{\boldsymbol{\epsilon}})$, the MCRB on the vector of the parameters of interest $\bar{\boldsymbol{\theta}}$ is given by:

$$\mathbf{MCRB}_{iid}(\bar{\boldsymbol{\theta}}) = \frac{\bar{\sigma}_n^2}{N}\mathbf{K}(\bar{\boldsymbol{\theta}})^{-1}. \tag{47}$$

It is worth highlighting here an interesting result: under the misspecificed scenario discussed in this section, i.e. when the data follow a CES, i.i.d. (true) model while the assumed one is a Gaussian, i.i.d., model, we have that:

$$\mathbf{A}(\bar{\boldsymbol{\theta}}) + \mathbf{B}_{iid}(\bar{\boldsymbol{\theta}}) = \mathbf{0}. \tag{48}$$

As explained in [6, Lemma 4.1], the result in (48), along with the block-diagonal structure of $\mathbf{MCRB}_{iid}(\bar{\boldsymbol{\epsilon}})$ in (46), implies that the *simplified Gaussian assumption does not lead to any degradation of the asymptotic estimation performance of the parameter vector of interest* $\bar{\boldsymbol{\theta}}$. In fact, $\mathbf{MCRB}_{iid}(\bar{\boldsymbol{\theta}})$ coincides with the lower bound that we can get if the true data model was an i.i.d. Gaussian one. This intriguing outcome can be explained through the semiparametric theory (refer to [17, Sec. IV;B] and [18, Sec. III.B]) that allows us to prove that the lack of knowledge of the density generator g does not have any asymptotic impact on the estimation of $\bar{\boldsymbol{\theta}}$.

Finally, if the true distribution is a Gaussian one, the term $E[\mathcal{Q}^2]$ is equal to 2 as proved in [9, Eq. (41)] and this lead us to the classical result about the CRB on the estimation of the variance in complex Gaussian data.

We note, in passing, that the same outcomes discussed in this section can be obtained, in a different yet equivalent way, from the Misspecified Slepian-Bangs formula [9], as discussed in [19].

6. An asymptotic efficient estimator under dependent observations

Let us go back now to the general misspecified nonlinear regression problem presented in Sec. 2. After having derived the MCRB for the vector of the parameters of interest $\bar{\boldsymbol{\theta}}$ in (36), the crucial question that arises is as follows: *is it possible to derive, under the misspecified Gaussian model* $\mathcal{F}_{\boldsymbol{\epsilon}}$ *in* (7)*, a consistent estimator* $\hat{\boldsymbol{\theta}}_N$ *of* $\bar{\boldsymbol{\theta}}$ *able to achieve the MCRB, at least asymptotically?*

It is well known that, under the i.i.d. case, the answer to this question is positive and $\hat{\boldsymbol{\theta}}_N$ is given by the Missmatched Maximum Likelihood estimator (MMLE) [1, 2, 7, 5, 6]. The extension to the dependent case has been provided in [4] where the asymptotic behaviour of the nonlinear least square estimator (NLLSE) for $\bar{\boldsymbol{\theta}}$ under the dependent data generating process in (2), that is:

$$\hat{\boldsymbol{\theta}}_N = \underset{\boldsymbol{\theta}\in\Theta}{\operatorname{argmin}} \left\{ \frac{1}{N} \sum\nolimits_{k=N_1}^{N_2} |x_k - f_k(\boldsymbol{\theta})|^2 \right\}, \tag{49}$$

has been investigated. Note that, when the misspecified Gaussian model $\mathcal{F}_{\boldsymbol{\epsilon}}$ in (7) is assumed, it is immediate to verify that, *for any finite* N, the NLLSE coincides with the MMLE. In fact, from (8), the misspecified log-likelihood function is $l(\boldsymbol{\theta}) = -N\ln(\pi\sigma_n^2) - ||\mathbf{x} - \mathbf{f}(\boldsymbol{\theta})||^2/\sigma_n^2$, then the MMLE for $\bar{\boldsymbol{\theta}}$ in $\mathcal{F}_{\boldsymbol{\epsilon}}$ is given by:

$$\underset{\boldsymbol{\theta}\in\Theta}{\operatorname{argmax}}\, l(\boldsymbol{\theta}) = \underset{\boldsymbol{\theta}\in\Theta}{\operatorname{argmin}} \left\{ ||\mathbf{x} - \mathbf{f}(\boldsymbol{\theta})||^2 \right\} = \hat{\boldsymbol{\theta}}_N, \tag{50}$$

that is the same estimator of the one in (49).

Remarkably, in [4] it has been proved that:

Theorem 2. *Under Assumption 1 and other technical regularity conditions (see A1-A9 in [4]),* $\hat{\boldsymbol{\theta}}_N$ *in* (49) *satisfies the following properties:*

1. *Consistency w.r.t. the true parameter vector:*

$$\hat{\boldsymbol{\theta}}_N \overset{a.s.}{\rightarrow} \bar{\boldsymbol{\theta}}, \tag{51}$$

where $\overset{a.s.}{\rightarrow}$ *indicates the almost sure convergence.*

2. *Asymptotic normality: Let us indicate as* $\underset{N\to\infty}{\sim}$ *the convergence in distribution, we have:*

$$\sqrt{N}\left[\mathbf{P}(\bar{\boldsymbol{\theta}})\right]^{-1/2}\mathbf{K}(\bar{\boldsymbol{\theta}})\left(\hat{\boldsymbol{\theta}}_N-\bar{\boldsymbol{\theta}}\right)\underset{N\to\infty}{\sim}\mathcal{N}(0,\mathbf{I}). \tag{52}$$

It follow directly from (52) that the asymptotic error covariance matrix of $\hat{\boldsymbol{\theta}}_N$ equates the $\mathbf{MCRB}(\bar{\boldsymbol{\theta}})$ in (36), i.e.:

$$\lim_{N\to\infty}\mathbf{w}^T\left(NE_{p_{\bar{\epsilon}}}\left[(\hat{\boldsymbol{\theta}}_N-\bar{\boldsymbol{\theta}})(\hat{\boldsymbol{\theta}}_N-\bar{\boldsymbol{\theta}})^\top\right]-\mathbf{K}(\bar{\boldsymbol{\theta}})^{-1}\mathbf{P}(\bar{\boldsymbol{\theta}})\mathbf{K}(\bar{\boldsymbol{\theta}})^{-1}\right)\mathbf{w}=0,\ \forall\mathbf{w}\in\mathbb{R}^P/\{\mathbf{0}\}. \tag{53}$$

Consequently, the NLLSE in (49) (that coincides with the MMLE under misspecified Gaussian assumption) is exactly the consistent and asymptotically efficient estimator that we were looking for.

To conclude, it can be noted that, in the i.i.d. case discussed in Sec. 5, eq. (53) simplifies to:

$$\lim_{N\to\infty}\mathbf{w}^T\left(NE_{p_{\bar{\epsilon}}}\left[(\hat{\boldsymbol{\theta}}_N-\bar{\boldsymbol{\theta}})(\hat{\boldsymbol{\theta}}_N-\bar{\boldsymbol{\theta}})^\top\right]-\bar{\sigma}_n^2\mathbf{K}(\bar{\boldsymbol{\theta}})^{-1}\right)\mathbf{w}=0,\ \forall\mathbf{w}\in\mathbb{R}^P/\{\mathbf{0}\}. \tag{54}$$

7. Application to Time-delay and Doppler estimation

Time-delay and Doppler estimation is fundamental in a plethora of engineering domains, including communications, radar, and navigation [20, 21, 22, 23, 24, 25, 26, 27, 28, 29], as it serves as the initial step at the receiver [24, 27, 28]. Due to its importance, understanding the achievable estimation performance in terms of MSE is of paramount practical interest. This crucial insight is typically provided by the CRB. Over the past few decades, numerous CRB expressions have been developed for time-delay and Doppler estimation problems, encompassing both finite narrow-band and wideband signals [21, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39]. Furthermore, recent studies have explored scenarios in which the actual signal model at the receiver differs from the assumed one [40, 41, 42, 43]. In these investigations, expressions for estimation bounds, as determined by the MCRB, have been established.

However all these prior studies share a common assumption: both the noise in the *true* signal model and the noise in the signal model *assumed* by the receiver follow a centered complex normal distribution with uncorrelated covariance matrix, i.e. a diagonal matrix. Surprisingly, despite the extensive research in this area, there is a notable absence in the literature regarding the ultimate attainable estimation performance for time-delay and Doppler (in terms of MSE) when the true signal model features a correlated non-Gaussian distributed noise. The aim of this section is then to fill this gap by relying on the theoretical results derived in the previous sections.

7.1. Signal model

We consider the transmitter T_X to receiver R_X direct transmission of a band-limited signal $a(t)$ with bandwidth B

$$a(t) = \sum_{n=N_1}^{N_2} a(nT_s)\ \operatorname{sinc}(\pi B (t - nT_s)),\ T_s = 1/B, \tag{55}$$

over a carrier with frequency f_c ($\lambda_c = c/f_c$, $\omega_c = 2\pi f_c$). The transmitter is located at position $\boldsymbol{P}_{T_X}(t)$ and the receiver is located at position $\boldsymbol{P}_{R_X}(t)$. The distance travelled by the transmitted signal is $\boldsymbol{P}_{T_X R_X} = \|\boldsymbol{P}_{T_X}(t - \tau_0(t)) - \boldsymbol{P}_{R_X}(t)\| \approx \frac{(\boldsymbol{P}_{T_X} - \boldsymbol{P}_{R_X})}{c} + \frac{v}{c}t$, that is, a first order approximation where $\bar{\tau} = \frac{(\boldsymbol{P}_{T_X} - \boldsymbol{P}_{R_X})}{c}$ and $\bar{b} = \frac{v}{c}$ with v the relative velocity between the transmitter and the receiver. Once the baseband demodulation process has been completed, the received signal can be expressed as [30, 36, 44]

$$x(t; \bar{\boldsymbol{\eta}}) = \bar{\alpha} a\left((t - \bar{\tau})(1 - \bar{b})\right) e^{-j2\pi f_c (\bar{b}(t - \bar{\tau}))} + n(t), \tag{56}$$

yielding to

$$x(t; \bar{\boldsymbol{\eta}}) = \bar{\alpha} a(t - \bar{\tau})\, e^{-j2\pi f_c (\bar{b}(t - \bar{\tau}))} + n(t), \tag{57}$$

under the narrowband assumption, i.e. the influence of the Doppler parameter on the baseband signal samples is omitted. The term $\bar{\alpha} = \bar{\rho} e^{j\bar{\Phi}}$ represents a complex gain, while $n(t)$ is a zero-mean, generally non Gaussian, wide sense stationary (WSS) continuous random process. The discrete signal model is built from $N = |N_1 - N_2 + 1|$ samples at $T_s = 1/F_s = 1/B$,

$$\mathbf{x} = \bar{\alpha}\boldsymbol{\mu}(\bar{\boldsymbol{\eta}}) + \mathbf{n} = \bar{\rho} e^{j\bar{\Phi}} \boldsymbol{\mu}(\bar{\boldsymbol{\eta}}) + \mathbf{n}, \tag{58}$$

with $\mathbf{x} = (\ldots, x(kT_s), \ldots)^\top$, $N_1 \leq k \leq N_2$ signal samples. Moreover, by posing $(\bar{\boldsymbol{\eta}}) = [\bar{\tau}, \bar{b}]^\top$, we have:

$$\boldsymbol{\mu}(\bar{\boldsymbol{\eta}}) = (\ldots, a(kT_s - \bar{\tau}) e^{-j2\pi f_c (\bar{b}(kT_s - \bar{\tau})}, \ldots)^\top. \tag{59}$$

Consequently, by defining the true vector of the parameters of interest as $\bar{\boldsymbol{\theta}}^\top = \left(\bar{\rho}, \bar{\Phi}, \bar{\boldsymbol{\eta}}^\top\right)$ the signal model in eq. (58) follows the form in (1):

$$\mathbf{x} = \bar{\alpha}\boldsymbol{\mu}(\bar{\boldsymbol{\eta}}) + \mathbf{n} = \mathbf{f}(\bar{\boldsymbol{\theta}}) + \mathbf{n}. \tag{60}$$

Finally, standard receivers assumes that the noise vector $\mathbf{n} \in \mathbb{C}^N$ is distributed as a centered complex normal random vector with diagonal covariance matrix, i.e. $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2 \boldsymbol{I}_N)$. Note that this represents the same misspecified scenario introduced in Sec. 2.2. Specifically, we have that the pdf of the observation vector $\mathbf{x}$ in (60) belongs to the misspecified model in (7), i.e. $\mathbf{x} \sim f_{\boldsymbol{\epsilon}} \in \mathcal{F}_{\boldsymbol{\epsilon}}$.

7.2. Time-delay and Doppler Closed-Form MCRB Expression for a Band-Limited Signal

It is interesting to note the likelihood between the expression obtained in previous sections and those already derived in the state of art. In particular, we may note that the matrix $-\mathbf{A}(\bar{\boldsymbol{\theta}})$ derived in (28) represents the FIM of a single source conditional signal model (CSM) [45]. A compact expression of this FIM, that depends only on the baseband signal samples, was recently derived in [36] as:

$$-\mathbf{A}(\bar{\boldsymbol{\theta}}) = \frac{2F_s}{\sigma_n^2} \mathrm{Re}\left\{\mathbf{Q}\mathbf{W}\mathbf{Q}^H\right\} \tag{61}$$

with

$$\mathbf{W} = \begin{bmatrix} w_1 & w_2^* & w_3^* \\ w_2 & W_{2,2} & w_4^* \\ w_3 & w_4 & W_{3,3} \end{bmatrix}, \tag{62a}$$

$$\mathbf{Q} = \begin{bmatrix} e^{j\bar{\Phi}} & 0 & 0 \\ j\bar{\alpha} & 0 & 0 \\ j\bar{\alpha}2\pi f_c \bar{b} & 0 & -\bar{\alpha} \\ 0 & -j\bar{\alpha}2\pi f_c & 0 \end{bmatrix}, \tag{62b}$$

where the elements of $\mathbf{W}$ can be expressed w.r.t. the baseband signal samples as,

$$\begin{aligned} w_1 &= \frac{1}{F_s}\mathbf{a}^H\mathbf{a}, \quad w_2 = \frac{1}{F_s^2}\mathbf{a}^H\mathbf{D}\mathbf{a}, \quad w_3 = \mathbf{a}^H\boldsymbol{\Lambda}\mathbf{a}, \\ w_4 &= \frac{1}{F_s}\mathbf{a}^H\mathbf{D}\boldsymbol{\Lambda}\mathbf{a}, \quad W_{2,2} = \frac{1}{F_s^3}\mathbf{a}^H\mathbf{D}^2\mathbf{a}, \quad W_{3,3} = F_s\mathbf{a}^H\mathbf{V}\mathbf{a}. \end{aligned} \tag{63}$$

with $\mathbf{a}$, the baseband samples vector, $\mathbf{D}$, $\boldsymbol{\Lambda}$ and $\mathbf{V}$ defined as,

$$\mathbf{a} = (\ldots, a(nT_s), \ldots)^\top_{N_1 \le n \le N_2}, \tag{64a}$$

$$\mathbf{D} = \mathrm{diag}(\ldots, n, \ldots)_{N_1 \le n \le N_2}, \tag{64b}$$

$$(\boldsymbol{\Lambda})_{n,n'} = \left| \begin{array}{l} n' \neq n : \frac{(-1)^{|n-n'|}}{n-n'} \\ n' = n : 0 \end{array} \right. \tag{64c}$$

$$(\mathbf{V})_{n,n'} = \left| \begin{array}{l} n' \neq n : (-1)^{|n-n'|} \frac{2}{(n-n')^2} \\ n' = n : \frac{\pi^2}{3} \end{array} \right. \tag{64d}$$

Moreover, under the uncorrelated noise assumption, we note from (46) that, since $\mathbf{B}_{iid}(\bar{\boldsymbol{\theta}}) = -\mathbf{A}(\bar{\boldsymbol{\theta}})$, the MCRB on the estimation of $\bar{\boldsymbol{\theta}}$ is given by:

$$\mathbf{MCRB}_{iid}(\bar{\boldsymbol{\theta}}) = \frac{2F_s}{\bar{\sigma}_n^2} \text{Re}\left\{\mathbf{QWQ}^H\right\} \tag{65}$$

345 In the case of correlated noise, the expression of the matrix $\mathbf{B}(\bar{\boldsymbol{\theta}}) = N^{-1}\bar{\sigma}_n^4\mathbf{P}(\bar{\boldsymbol{\theta}})$ in (26) is more challenging since $\mathbf{P}(\bar{\boldsymbol{\theta}})$ involves the autocorrelation function $r_n[j]$. We can distinguish between the following two cases. If $r_n[j]$ is *a-priori known*, for the application at hand $\mathbf{P}(\bar{\boldsymbol{\theta}})$ can be expressed as:

$$\mathbf{P}(\bar{\boldsymbol{\theta}}) = \frac{2F_s\bar{\sigma}_n^2}{N} \text{Re}\left\{\mathbf{QWQ}^H\right\} + \frac{4F_s}{N}\sum_{j=1}^{l} \text{Re}\left\{r_n[j]\mathbf{QW}_j\mathbf{Q}^H\right\} \tag{66}$$

where l is the correlation lag defined in Theo. 1 and:

$$\mathbf{W}_j = \begin{bmatrix} w_1^{(j)} & w_2^{(j)*} & w_3^{(j)*} \\ w_2^{(j)} & W_{2,2}^{(j)} & w_4^{(j)*} \\ w_3^{(j)} & w_4^{(j)} & W_{3,3}^{(j)} \end{bmatrix}, \tag{67}$$

and the elements of $\mathbf{W}_j$ can be expressed w.r.t. the baseband signal samples as,

$$\begin{aligned} w_1^{(j)} &= \frac{1}{F_s}\mathbf{a}_{j^+}^H\mathbf{a}_{j^-}, \quad w_2^{(j)} = \frac{1}{F_s^2}\mathbf{a}_{j^+}^H\mathbf{D}_j\mathbf{a}_{j^-}, \\ w_3^{(j)} &= \mathbf{a}_{j^+}^H\boldsymbol{\Lambda}_j\mathbf{a}_{j^-}, \quad w_4^{(j)} = \frac{1}{F_s}\mathbf{a}_{j^+}^H\mathbf{D}_j\boldsymbol{\Lambda}_j\mathbf{a}_{j^-}, \\ W_{2,2}^{(j)} &= \frac{1}{F_s^3}\mathbf{a}_{j^+}^H\mathbf{D}_j^2\mathbf{a}_{j^-}, \quad W_{3,3}^{(j)} = F_s\mathbf{a}_{j^+}^H\mathbf{V}_j\mathbf{a}_{j^-}. \end{aligned} \tag{68}$$

350 with $\mathbf{a}_{j^-}$, $\mathbf{a}_{j^+}$, $\mathbf{D}_j$, $\boldsymbol{\Lambda}_j$ and $\mathbf{V}_j$ defined as,

$$\mathbf{a}_{j^+} = (\ldots, a(nT_s), \ldots)^\top_{N_1+j\leq n\leq N_2}, \tag{69a}$$

$$\mathbf{a}_{j^-} = (\ldots, a(nT_s), \ldots)^\top_{N_1\leq n\leq N_2-j}, \tag{69b}$$

$$\mathbf{D}_j = \text{diag}\left(\ldots, n, \ldots\right)_{N_1\leq n\leq N_2-j}, \tag{69c}$$

$$\left(\boldsymbol{\Lambda}_j\right)_{n,n'} = \left| \begin{array}{l} n' \neq n : \frac{(-1)^{|n-n'|}}{n-n'} \\ n' = n : 0 \end{array} \right. \tag{69d}$$

$$\left(\mathbf{V}_j\right)_{n,n'} = \left| \begin{array}{l} n' \neq n : (-1)^{|n-n'|} \frac{2}{(n-n')^2} \\ n' = n : \frac{\pi^2}{3} \end{array} \right. \tag{69e}$$

with $N_1 \leq n, n' \leq N_2 - j$. If $r_n[j]$ is *not a-priori known*, from (37), a consistent estimator of $\mathbf{P}(\bar{\boldsymbol{\theta}})$ can be implemented as:

$$\begin{aligned} \widehat{\mathbf{P}}_N = N^{-1} 2F_s \mathrm{Re}\left\{\mathbf{Q}\mathbf{W}\mathbf{Q}^H\right\} \sum\nolimits_{k=N_1}^{N_2} |\hat{n}_k|^2 + \\ + N^{-1} 4F_s \sum\nolimits_{j=1}^{l} \mathrm{Re}\left\{\hat{n}_{k+j}^* \hat{n}_k \mathbf{Q}\mathbf{W}_j\mathbf{Q}^H\right\}, \end{aligned} \tag{70}$$

where $\hat{n}_k \triangleq x_k - f_k(\hat{\boldsymbol{\theta}}_N)$ and $\hat{\boldsymbol{\theta}}_N$ is the consistent estimator of $\bar{\boldsymbol{\theta}}$ defined as [43] [5]:

$$\hat{\boldsymbol{\eta}} = \arg\max_{\boldsymbol{\eta}} \left\|\mathbf{\Pi}_{\mu(\boldsymbol{\eta})}\mathbf{x}\right\|^2 \tag{71}$$

$$\widehat{\rho} = \left|\left[\boldsymbol{\mu}^H(\hat{\boldsymbol{\eta}})\,\boldsymbol{\mu}(\hat{\boldsymbol{\eta}})\right]^{-1} \boldsymbol{\mu}^H(\hat{\boldsymbol{\eta}})\,\mathbf{x}\right| \tag{72}$$

$$\widehat{\Phi} = \arg\left\{\left[\boldsymbol{\mu}^H(\hat{\boldsymbol{\eta}})\,\boldsymbol{\mu}(\hat{\boldsymbol{\eta}})\right]^{-1} \boldsymbol{\mu}^H(\hat{\boldsymbol{\eta}})\,\mathbf{x}\right\} \tag{73}$$

8. Simulation and Discussion

To support our theoretical analysis, we examine the transmission and reception of a GPS L1 C/A signal [29]. This signal employs a baseband signal represented by a periodic binary phase-shift keying (BPSK) Gold code with a length of 1023 chips of period 1ms. At the receiver, we set a sampling frequency $F_s = 4$ MHz, which is the standard rate for most commercial receivers. The GNSS receiver assumes that the noise follows a standard centered normal distribution.

Scenario 1 In a first scenario, we set a true signal model where the noise is sampled from an autoregressive complex discrete random process with innovations following a complex centered t-distribution [6, Sec. 4.6.1.1] with $\upsilon > 1$ degrees of freedom (or *shape parameter*) that control the level of non-Gaussianity and a scale parameter μ. The second-order modular variate Q of a t-distribution is F-distributed according to [6, eq. 4.59]. Then, in order to meet the constraint $E[Q] = 1$, the scale as to be set as $\mu = \frac{\upsilon}{\bar{\sigma}_n^2(\upsilon-1)}$ (see [6, eq. 4.60]) where $\bar{\sigma}_n^2$ depends on the signal to noise ratio at the output of the match filter SNR_{out}. The SNR_{out} is defined as:

$$SNR_{out} = \frac{|\alpha|^2 \mathbf{a}^H \mathbf{a}}{\bar{\sigma}_n^2}. \tag{74}$$

Furthermore, in this scenario, we employ two autoregressive processes (AR) of order 1 and 6, respectively, to model the noise correlation. The poles of the process are set to $p = 0.9 \cdot e^{j2\pi \cdot 0.3}$ for the order 1 process

[5] Let $S = span(\mathbf{A})$, with $\mathbf{A}$ a matrix, be the linear span of the set of its column vectors. The orthogonal projector over S is $\mathbf{\Pi}_{\mathbf{A}} = \mathbf{A}\left(\mathbf{A}^H\mathbf{A}\right)\mathbf{A}^H$.

and $\boldsymbol{p} = [0.5 \cdot e^{-j2\pi \cdot 0.4}, 0.6 \cdot e^{-j2\pi \cdot 0.2}, 0.7, 0.4 \cdot e^{j2\pi \cdot 0.1}, 0.5 \cdot e^{j2\pi \cdot 0.3}, 0.6 \cdot e^{j2\pi \cdot 0.35}]$ for the order 6 process. The MMLE for the joint estimation of the time-delay and Doppler is defined in (71). The root mean square error (RMSE) results of the MMLE for the parameters of interest $\boldsymbol{\eta}^T = [\tau, b]$ are shown in Figs. 1 and 2 w.r.t. the SNR_{out}. The number of Monte Carlo is set to 1000 iterations and $\upsilon = 2.5$. In the results one can observe that the RMSE ($\sqrt{MSE}$) of the pseudotrue parameter converges to the asymptotic estimation performance derived in Sec. 4. These results confirm the theoretical derivation. Moreover, the Gaussian i.d.d $\sqrt{CRB}$ has been included to quantify the performance with respect to the correlated case. It is worth to underline that the previous theoretical results are valid for any joint pdf $p_{\mathbf{n}}$ admitting finite first and second order moments and not only of the one obtained from a t-distribution.

Scenario 2 In the second scenario, we would like to illustrate the estimation performance in the case where the noise is i.i.d. To do so, we propose to set the true signal model as in the previous scenario, i.e the noise process is complex centered t-distributed (we remind that in this particular case, the noise is assumed to be i.i.d and there is not need to define any process to characterized the noise correlation). Moreover, we also set a true signal model where the noise is distributed according to a complex centered Generalized Gaussian (GG) distribution, [6, Sec. 4.6.1.2] with exponent $s > 0$ and scale $b > 0$, where s is a parameter controlling the level of non-Gaussianity. The second-order modular variate $\mathcal{Q}$ of a GG distribution is given by $\mathcal{Q} =_d G^{1/s}$ where G is a Gamma distributed random variable with parameter $1/s$ and b, i.e. $G \sim \text{Gam}(1/s, b)$ [15, Sec. IV.B]. In order to satisfy the constraint $E\{\mathcal{Q}\} = 1$ (see Sec. 4), we set $b = \left(\frac{\bar{\sigma}_n^2 \Gamma(1/s)}{\Gamma(2/s)}\right)^s$. Again, $\bar{\sigma}_n^2$ set the SNR_{out}. The RMSE results of the MMLE for the parameters of interest $\boldsymbol{\eta}^T = [\tau, b]$ are shown in Figs. 3 and 4 w.r.t. the SNR_{out}. The number of Monte Carlo is set to 1000 iterations. In the simulation, complex centered Generalized Gaussian distributions with $s = \{0.5, 1.5, 2.5\}$ and complex centered t-distributions with $\upsilon = \{1.1, 2, 3\}$ have been used as a true model. In the results one can observe that the RMSE of the pseudotrue parameter converges to the asymptotic estimation performance derived in Sec. 4. These results confirm the theoretical derivation. Note also that the $\sqrt{MCRB}$ is equal to the $\sqrt{CRB}$. It is important to underline that the preceding theoretical findings are applicable to all true noise models characterized by Complex Elliptically Symmetric (CES) distributions, not limited to the Gaussian (GG) and t-distribution cases. As mentioned earlier, a thorough explanation of this phenomenon is rooted in semiparametric theory (refer to [17, Sec. IV.B] and [18, Sec. III.B]), and a comprehensive explanation will be provided in future research. For now, we restrict our discussion to this observation: the equivalence between the MCRB and the CRB holds true only when the parameters of interest parameterize the mean of the observation vectors. Conversely, if some parameter of interest is involved in the covariance matrix of the observations, this equivalence may no longer be valid.

9. Conclusion

This paper focused on the performance evaluation of estimation procedures in nonlinear regression models. In particular, we were interested in analyzing the asymptotic performance of inference algorithms based on the simplistic i.i.d. Gaussian assumption in the presence of correlated and non-Gaussian noise. To this end, the related MCRB has been evaluated and the consistency and efficiency properties of the MMLE/NLLSE investigated. Under a weak condition on the rate of decay of the autocorrelation function (Assumption 1), our results show that:

- The MCRB for the parameters of interest in the nonlinear regression model depends of the autocorrelation function of the noise but not on the joint pdf on the noise samples that can then be left unspecified. Moreover, the MMLE/NLLSE is consistent and efficient with respect to the relevant MCRB.

- If the noise samples are modeled as zero-mean, i.i.d. CES-distributed (with unspecified density generator) random variables, the MCRB on the parameter of interest equates the CRB derived under i.i.d. Gaussian assumption. This means that the asymptotic performance of Gaussian-based MMLE, i.e. the NLLSE, are not affected by the lack of knowledge of the true non-Gaussian and heavy-tailed noise distribution.

Since the i.i.d. Gaussian assumption is widely used in applications, these theoretical results are of great practical interest. Specifically, they implies that, a practitioner can continue to use the i.i.d. Gaussian-based inference procedures without any loss in asymptotic estimation performance even when the noise samples are heavy-tailed, non-Gaussian but still i.i.d. random variables. On the other hands, if the noise samples are correlated heavy-tailed, non-Gaussian random variables, the asymptotic performance of i.i.d. Gaussian-based procedures depends only on the autocorrelation function of the noise process and not on the specific joint pdf of its samples. Our theoretical findings have been then used to investigate the asymptotic performance of Gaussian procedure in time-delay and Doppler estimation for GNSS.

Future works will focus on the possibility to derive lower bounds on the performance of estimation procedures in the presence of an non-perfect knowledge of the nonlinear function characterizing the regression model.

Acknowledgments

This work was supported by TéSA.

Appendix

In this Appensix, we provide a simple proof that, in equation (12), the minimization w.r.t. to $\boldsymbol{\theta}$ is independent from the one of σ_n^2. Under some differentiability conditions, we have that:

$$
\begin{aligned}
&\arg\min_{\boldsymbol{\epsilon}\in\Gamma}\left\{E_{p_{\bar{\epsilon}}}\left[\frac{1}{\sigma_n^2}\left[\|\mathbf{x}-\mathbf{f}(\bar{\boldsymbol{\theta}})\|^2\right]\right]+N\ln(\sigma_n^2)\right\}\\
&\Rightarrow \nabla_{\boldsymbol{\epsilon}}\left(E_{p_{\bar{\epsilon}}}\left[\frac{1}{\sigma_n^2}\left[\|\mathbf{x}-\mathbf{f}(\bar{\boldsymbol{\theta}})\|^2\right]\right]+N\ln(\sigma_n^2)\right)|_{\boldsymbol{\epsilon}=\boldsymbol{\epsilon}_0}=\mathbf{0}\\
&\Rightarrow \begin{cases}\nabla_{\boldsymbol{\theta}}\left(E_{p_{\bar{\epsilon}}}\left[\frac{1}{\sigma_n^2}\left[\|\mathbf{x}-\mathbf{f}(\bar{\boldsymbol{\theta}})\|^2\right]\right]+N\ln(\sigma_n^2)\right)|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}=\mathbf{0}\\ \nabla_{\sigma_n^2}\left(E_{p_{\bar{\epsilon}}}\left[\frac{1}{\sigma_n^2}\left[\|\mathbf{x}-\mathbf{f}(\bar{\boldsymbol{\theta}})\|^2\right]\right]+N\ln(\sigma_n^2)\right)|_{\sigma_n^2=\sigma_0^2}=\mathbf{0}\end{cases}
\end{aligned}
\tag{75}
$$

Now, it is immediate to verify that the first equation of the non-linear system associated to the gradient can be expressed as:

$$
\begin{aligned}
&\nabla_{\boldsymbol{\theta}}\left(E_{p_{\bar{\epsilon}}}\left[\frac{1}{\sigma_n^2}\left[\|\mathbf{x}-\mathbf{f}(\bar{\boldsymbol{\theta}})\|^2\right]\right]+N\ln(\sigma_n^2)\right)|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}=\mathbf{0}\\
&\Rightarrow \frac{1}{\sigma_n^2}\left(\nabla_{\boldsymbol{\theta}}E_{p_{\bar{\epsilon}}}\left[\left[\|\mathbf{x}-\mathbf{f}(\bar{\boldsymbol{\theta}})\|^2\right]\right]+\nabla_{\boldsymbol{\theta}}N\ln(\sigma_n^2)\right)|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}=\mathbf{0}\\
&\Rightarrow \nabla_{\boldsymbol{\theta}}E_{p_{\bar{\epsilon}}}\left[\|\mathbf{x}-\mathbf{f}(\bar{\boldsymbol{\theta}})\|^2\right]|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}=\mathbf{0},
\end{aligned}
\tag{76}
$$

that does not depend on σ_n^2. To conclude the proof, we can immediately note that this equation is equal to (13b).

References

[1] P. J. Huber, The behavior of maximum likelihood estimates under nonstandard conditions, in: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics, University of California Press, Berkeley, Calif., 1967, pp. 221–233.

[2] H. White, Maximum likelihood estimation of misspecified models, Econometrica: Journal of the Econometric Society (1982) 1 – 25.

[3] Q. H. Vuong, Cramér-Rao bounds for misspecified models, California Inst. of Technol., Pasadena, CA, USA, Working Paper 652, Oct. 1986 [Online]. Available: https://www.hss.caltech.edu/content/cramer-rao-bounds-misspeci- fied-models.

[4] H. White, I. Domowitz, Nonlinear regression with dependent observations, Econometrica 52 (1) (1984) 143–161.

[5] S. Fortunati, F. Gini, M. S. Greco, C. D. Richmond, Performance bounds for parameter estimation under misspecified models: Fundamental findings and applications, IEEE Signal Processing Magazine 34 (6) (2017) 142–157.

[6] S. Fortunati, F. Gini, M. S. Greco, Chapter 4 - Parameter bounds under misspecified models for adaptive radar detection, in: R. Chellappa, S. Theodoridis (Eds.), Academic Press Library in Signal Processing, Volume 7, Academic Press, 2018, pp. 197–252.

[7] C. D. Richmond, L. L. Horowitz, Parameter bounds on estimation accuracy under model misspecification, IEEE Trans. Signal Process. 63 (9) (2015) 2263–2278. `doi:https://doi.org/10.1109/TSP.2015.2411222`.

[8] S. Fortunati, F. Gini, M. S. Greco, The constrained Misspecified Cramér-Rao bound, IEEE Signal Processing Letters 23 (5) (2016) 718–721.

[9] A. Mennad, S. Fortunati, M. N. El Korso, A. Younsi, A. Zoubir, A. Renaux, Slepian-bangs-type formulas and the related misspecified cramér-rao bounds for complex elliptically symmetric distributions, Signal Processing 142 (2018) 320–329.

[10] K. Kreutz-Delgado, The complex gradient operator and the CR-calculus, in: ISI World Statistics Congress 2017 (ISI2017), 2017.
URL `https://arxiv.org/abs/0906.4835`

[11] S. M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory, Prentice-Hall, Englewood Cliffs, New Jersey, USA, 1993.

[12] P. Stoica, R. L. Moses, Spectral Analysis of Signals, 2nd Edition, Pearson Prentice Hall, 2005.

[13] B. Picinbono, On circularity, IEEE Transactions on Signal Processing 42 (12) (1994) 3473–3482.

[14] S. Fortunati, L. Sanguinetti, F. Gini, M. S. Greco, B. Himed, Massive mimo radar for target detection, IEEE Transactions on Signal Processing 68 (2020) 859–871.

[15] E. Ollila, D. E. Tyler, V. Koivunen, H. V. Poor, Complex elliptically symmetric distributions: Survey, new results and applications, IEEE Transactions on Signal Processing 60 (11) (2012) 5597–5625. doi:10.1109/TSP.2012.2212433.

[16] A. W. van der Vaart, Asymptotic Statistics, Cambridge University Press, 1998.

[17] S. Fortunati, F. Gini, M. S. Greco, A. M. Zoubir, M. Rangaswamy, Semiparametric inference and lower bounds for real elliptically symmetric distributions, IEEE Transactions on Signal Processing 67 (1) (2019) 164–177.

[18] S. Fortunati, F. Gini, M. S. Greco, A. M. Zoubir, M. Rangaswamy, Semiparametric CRB and Slepian-Bangs formulas for complex elliptically symmetric distributions, IEEE Transactions on Signal Processing 67 (20) (2019) 5352–5364. doi:10.1109/TSP.2019.2939084.

[19] L. Ortega, F. Fortunati, Misspecified time-delay and doppler estimation over non-gaussian scenarios, accepted in ICASSP 2024.

[20] D. A. Swick, A Review of Wideband Ambiguity Functions, Tech. Rep. 6994, Naval Res. Lab., Washington DC (1969).

[21] H. L. Van Trees, Detection, Estimation, and Modulation Theory, Part III: Radar – Sonar Signal Processing and Gaussian Signals in Noise, J. Wiley & Sons, 2001.

[22] U. Mengali, A. N. D'Andrea, Synchronization Techniques for Digital Receivers, Plenum Press, New York, USA, 1997.

[23] H. L. Van Trees, Optimum Array Processing, Wiley-Interscience, New-York, 2002.

[24] D. W. Ricker, Echo Signal Processing, Kluwer Academic, Springer, New York, USA, 2003.

[25] J. Chen, Y. Huang, J. Benesty, Time delay estimation, in: Y. Huang, J. Benesty (Eds.), Audio Signal Processing for Next-Generation Multimedia Communication Systems, Springer, Boston, MA, USA, 2004, Ch. 8, pp. 197–227.

[26] B. C. Levy, Principles of Signal Detection and Parameter Estimation, Springer, 2008.

[27] F. B. D. Munoz, C. Vargas, R. Enriquez, Position Location Techniques and Applications, Academic Press, Oxford, GB, 2009.

[28] J. Yan, et al., Review of range-based positioning algorithms, IEEE Trans. Aerosp. Electron. Syst. 28 (8) (2013) 2–27.

[29] P. J. G. Teunissen, O. Montenbruck (Eds.), Handbook of Global Navigation Satellite Systems, Springer, Switzerland, 2017. doi:https://doi.org/10.1007/978-3-319-42928-1.

[30] A. Dogandzic, A. Nehorai, Cramér-Rao bounds for estimating range, velocity, and direction with an active array, IEEE Trans. Signal Process. 49 (6) (2001) 1122–1137. doi:https://doi.org/10.1109/SAM.2000.878032.

[31] N. Noels, H. Wymeersch, H. Steendam, M. Moeneclaey, True Cramér-Rao bound for timing recovery from a bandlimited linearly modulated waveform with unknown carrier phase and frequency, IEEE Trans. on Communications 52 (3) (2004) 473–483.

[32] T. Zhao, T. Huang, Cramér-Rao lower bounds for the joint delay-doppler estimation of an extended target, IEEE Trans.

Signal Process. 64 (6) (2016) 1562–1573.

[33] Y. Chen, R. S. Blum, On the Impact of Unknown Signals on Delay, Doppler, Amplitude, and Phase Parameter Estimation, IEEE Trans. Signal Process. 67 (2) (2019) 431–443.

[34] Q. Jin, K. M. Wong, Z.-Q. Luo, The Estimation of Time Delay and Doppler Stretch of Wideband Signals, IEEE Trans. Signal Process. 43 (4) (1995) 904–916.

[35] X. Niu, P. C. Ching, Y. Chan, Wavelet Based Approach for Joint Time Delay and Doppler Stretch Measurements, IEEE Trans. Aerosp. Electron. Syst. 35 (3) (1999) 1111–1119.

[36] D. Medina, L. Ortega, J. Vilà-Valls, P. Closas, F. Vincent, E. Chaumette, Compact CRB for delay, doppler and phase estimation - application to GNSS SPP & RTK performance characterization, IET Radar, Sonar & Navigation 14 (10) (2020) 1537–1549. doi:https://doi.org/10.1049/iet-rsn.2020.0168.

[37] P. Das, J. Vilà-Valls, F. Vincent, L. Davain, E. Chaumette, A New Compact Delay, Doppler Stretch and Phase Estimation CRB with a Band-Limited Signal for GenE. Remote Sensing Applications, Remote Sensing 12 (18) (2020) 2913.

[38] C. Lubeigt, L. Ortega, J. Vilà-Valls, L. Lestarquit, E. Chaumette, Joint delay-doppler estimation performance in a dual source context, Remote Sensing 12 (23). doi:https://doi.org/10.3390/rs12233894.

[39] H. McPhee, L. Ortega, J. Vilà-Valls, E. Chaumette, Accounting for acceleration—signal parameters estimation performance limits in high dynamics applications, IEEE Transactions on Aerospace and Electronic Systems 59 (1) (2023) 610–622. doi:https://doi.org/10.1109/TAES.2022.3189611.

[40] C. Lubeigt, L. Ortega, J. Vilà-Valls, E. Chaumette, Untangling first and second order statistics contributions in multipath scenarios, Signal Processing 205 (2023) 108868. doi:https://doi.org/10.1016/j.sigpro.2022.108868.

[41] H. McPhee, L. Ortega, J. Vilà-Valls, E. Chaumette, On the accuracy limits of misspecified delay-doppler estimation, Signal Processing 205 (2023) 108872. doi:https://doi.org/10.1016/j.sigpro.2022.108872.

[42] L. Ortega, J. Vilà-Valls, E. Chaumette, Theoretical evaluation of the GNSS synchronization performance degradation under interferences, Denver, CO, USA, 2022. doi:https://doi.org/10.33012/2022.18564.

[43] L. Ortega, C. Lubeigt, J. Vilà-Valls, E. Chaumette, On gnss synchronization performance degradation under interference scenarios: Bias and misspecified cramér-rao bounds, NAVIGATION: Journal of the Institute of Navigation 70 (4).

[44] M. I. Skolnik, Radar Handbook, 3rd Edition, McGraw-Hill, New York, USA, 1990. doi:https://doi.org/10.1109/MAES.2008.4523916.

[45] P. Stoica, A. Nehorai, Performances study of conditional and unconditional direction of arrival estimation, IEEE Trans. Acoust., Speech, Signal Process. 38 (10) (1990) 1783–1795. doi:https://doi.org/10.1109/29.60109.

$$E_{p_{\epsilon}}\left[\nabla_{\boldsymbol{\theta}} \ln f_{\boldsymbol{\theta}}(\mathbf{x};\bar{\boldsymbol{\epsilon}}) \nabla_{\boldsymbol{\theta}}^{\top} \ln f_{\boldsymbol{\epsilon}}(\mathbf{x};\bar{\boldsymbol{\epsilon}})\right] = E_{p_{\mathrm{n}}}\left[\left[\frac{2}{\bar{\sigma}_n^2}\sum_{k=N_1}^{N_2} \mathrm{Re}\left\{n_k^* \nabla_{\boldsymbol{\theta}} \bar{f}_k\right\}\right]\left[\frac{2}{\bar{\sigma}_n^2}\sum_{j=N_1}^{N_2} \mathrm{Re}\left\{n_j^* \nabla_{\boldsymbol{\theta}} \bar{f}_j\right\}\right]^T\right] \tag{26a}$$

$$= \frac{4}{\bar{\sigma}_n^4}\sum_{k=N_1}^{N_2}\sum_{j=N_1}^{N_2} E_{p_{\mathrm{n}}}\left[\mathrm{Re}\left\{n_k^* \nabla_{\boldsymbol{\theta}} \bar{f}_k\right\} \mathrm{Re}\left\{n_j^* \nabla_{\boldsymbol{\theta}}^T \bar{f}_j\right\}\right] = \tag{26b}$$

$$\frac{4}{\bar{\sigma}_n^4}\sum_{k=N_1}^{N_2}\sum_{j=N_1}^{N_2} E_{p_{\mathrm{n}}}\left[\left(\frac{n_k^* \nabla_{\boldsymbol{\theta}} \bar{f}_k + n_k \nabla_{\boldsymbol{\theta}}^* \bar{f}_k}{2}\right)\left(\frac{n_j^* \nabla_{\boldsymbol{\theta}}^T \bar{f}_j + n_j \nabla_{\boldsymbol{\theta}}^H \bar{f}_j}{2}\right)\right] \tag{26c}$$

$$= \frac{2}{\bar{\sigma}_n^4}\sum_{k=N_1}^{N_2}\sum_{j=N_1}^{N_2}\left[\mathrm{Re}\left\{\left(E_{p_{\mathrm{n}}}\left[n_k n_j\right]\right)^* \nabla_{\boldsymbol{\theta}} \bar{f}_k \nabla_{\boldsymbol{\theta}}^T \bar{f}_j\right\} + \mathrm{Re}\left\{E_{p_{\mathrm{n}}}\left[n_k^* n_j\right] \nabla_{\boldsymbol{\theta}} \bar{f}_k \nabla_{\boldsymbol{\theta}}^H \bar{f}_j\right\}\right] \tag{26d}$$

$$= \frac{2}{\bar{\sigma}_n^4}\sum_{k=N_1}^{N_2} E_{p_{\mathrm{n}}}\left[|n_k|^2\right] \mathrm{Re}\left\{\nabla_{\boldsymbol{\theta}} \bar{f}_k \nabla_{\boldsymbol{\theta}}^H \bar{f}_k\right\} + \frac{2}{\bar{\sigma}_n^4}\sum_{k=N_1}^{N_2}\sum_{\substack{j=N_1\\ j\neq k}}^{N_2} \mathrm{Re}\left\{E_{p_{\mathrm{n}}}\left[n_k^* n_j\right] \nabla_{\boldsymbol{\theta}} \bar{f}_k \nabla_{\boldsymbol{\theta}}^H \bar{f}_j\right\} \tag{26e}$$

$$= \frac{2}{\bar{\sigma}_n^4}\sum_{k=N_1}^{N_2} E_{p_{\mathrm{n}}}\left[|n_k|^2\right] \mathrm{Re}\left\{\nabla_{\boldsymbol{\theta}} \bar{f}_k \nabla_{\boldsymbol{\theta}}^H \bar{f}_k\right\} + \frac{4}{\bar{\sigma}_n^4}\sum_{j=1}^{N_2-N_1}\sum_{k=N_1}^{N_2-j} \mathrm{Re}\left\{E_{p_{\mathrm{n}}}\left[n_{k+j}^* n_k\right] \nabla_{\boldsymbol{\theta}} \bar{f}_{k+j} \nabla_{\boldsymbol{\theta}}^H \bar{f}_k\right\} \tag{26f}$$

$$= \frac{N}{\bar{\sigma}_n^4}\left[\frac{2}{N}\sum_{k=N_1}^{N_2} r_n[0]\mathrm{Re}\left\{\nabla_{\boldsymbol{\theta}} \bar{f}_k \nabla_{\boldsymbol{\theta}}^H \bar{f}_k\right\} + \frac{4}{N}\sum_{j=1}^{N_2-N_1}\sum_{k=N_1}^{N_2-j} \mathrm{Re}\left\{r_n[j] \nabla_{\boldsymbol{\theta}} \bar{f}_{k+j} \nabla_{\boldsymbol{\theta}}^H \bar{f}_k\right\}\right] \triangleq \frac{N}{\bar{\sigma}_n^4}\mathbf{P}(\bar{\boldsymbol{\theta}}). \tag{26g}$$

$$E_{p_{\mathrm{n}}}\left[(\mathbf{n}^H \mathbf{n})^2\right] = \bar{\sigma}_n^4 E\left[\left(\sum_{i=N_1}^{N_2} Q_i |u_i|^2\right)^2\right] = \bar{\sigma}_n^4 E\left[\sum_{i=N_1}^{N_2}\sum_{j=N_1}^{N_2} Q_i Q_j |u_i|^2 |u_j|^2\right] = \tag{42a}$$

$$\bar{\sigma}_n^4 E\left[\sum_{i=N_1}^{N_2} Q_i^2 |u_i|^4\right] + \bar{\sigma}_n^4 E\left[\sum_{i=N_1}^{N_2}\sum_{j=N_1, j\neq i}^{N_2} Q_i Q_j |u_i|^2 |u_j|^2\right] = \tag{42b}$$

$$\bar{\sigma}_n^4\left[\sum_{i=N_1}^{N_2} E\left[Q_i^2\right] E\left[|u_i|^4\right] + \sum_{i=N_1}^{N_2}\sum_{j=N_1, j\neq i}^{N_2} E\left[Q_i\right] E\left[Q_j\right] E\left[|u_i|^2\right] E\left[|u_j|^2\right]\right] \tag{42c}$$

$$= \bar{\sigma}_n^4\left(N E\left[Q^2\right] + N(N-1)\right) = \bar{\sigma}_n^4\left(N(E\left[Q^2\right] - 1) + N^2\right), \tag{42d}$$

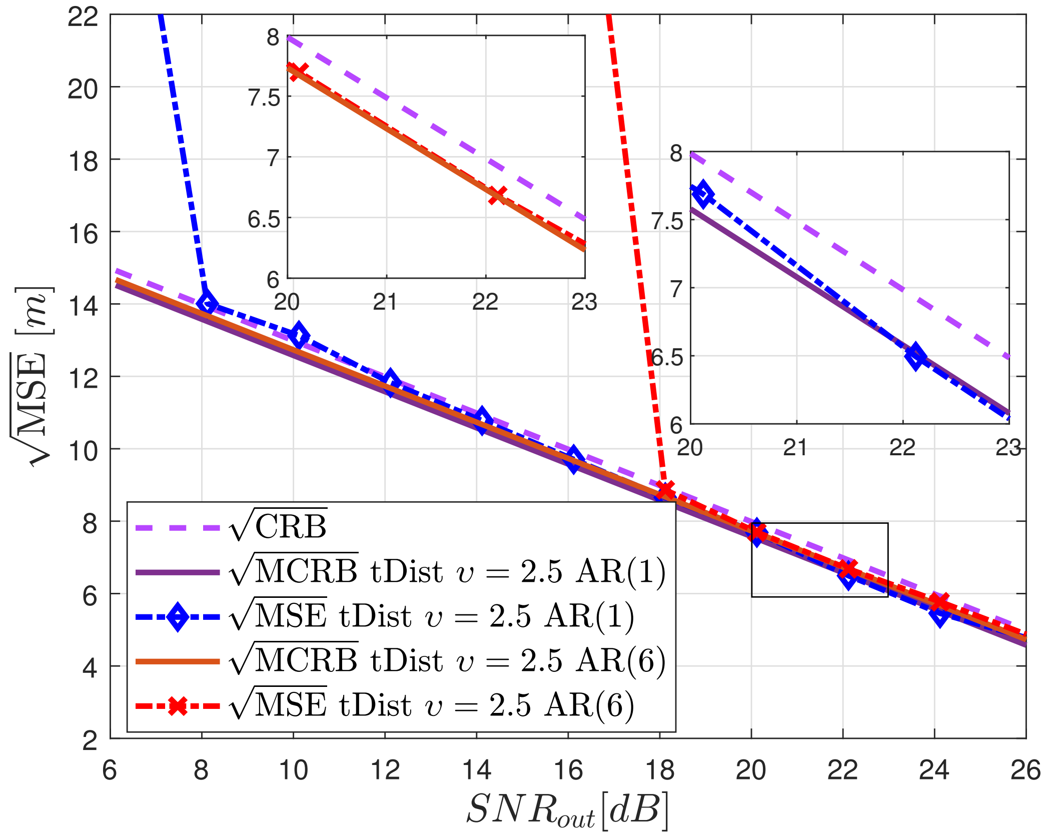

Figure 1: RMSE of the MMLE of the time-delay considering complex centered t-dist. with $\upsilon = 2.5$ and two AR processes of order 1 and 6, respectively, to model the noise correlation.

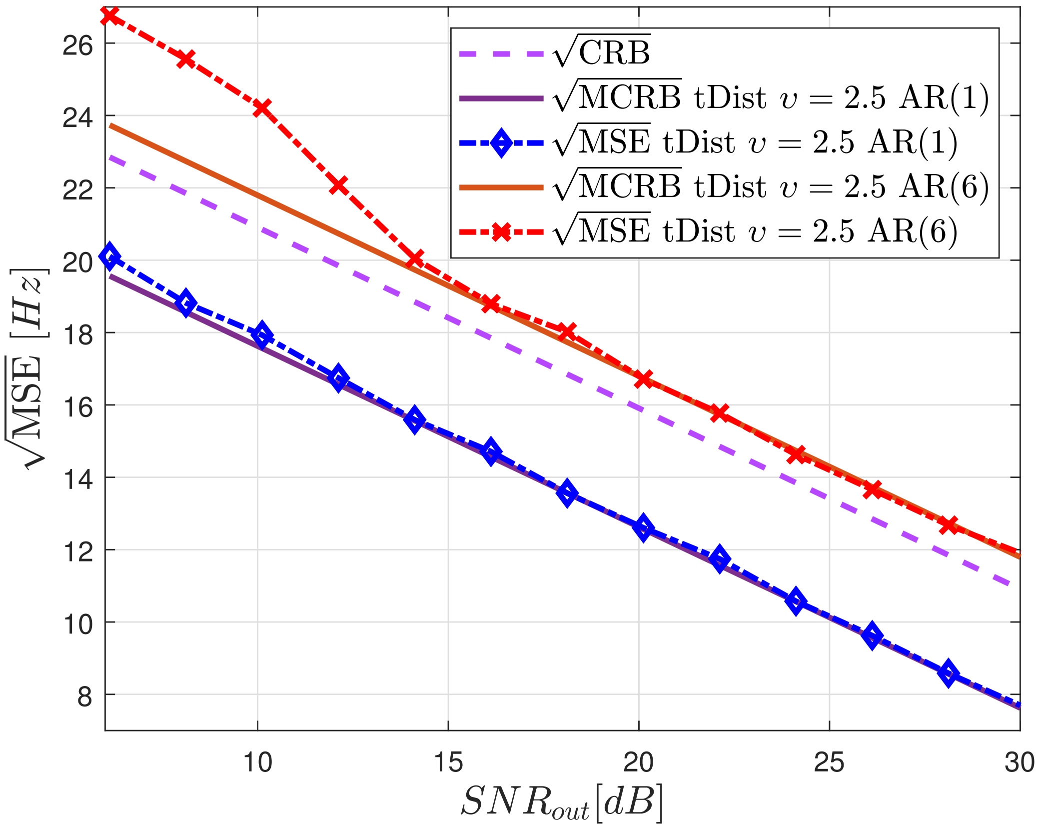

Figure 2: RMSE of the MMLE of the Doppler considering complex centered t-dist. with $\upsilon = 2.5$ and two AR processes of order 1 and 6, respectively, to model the noise correlation.

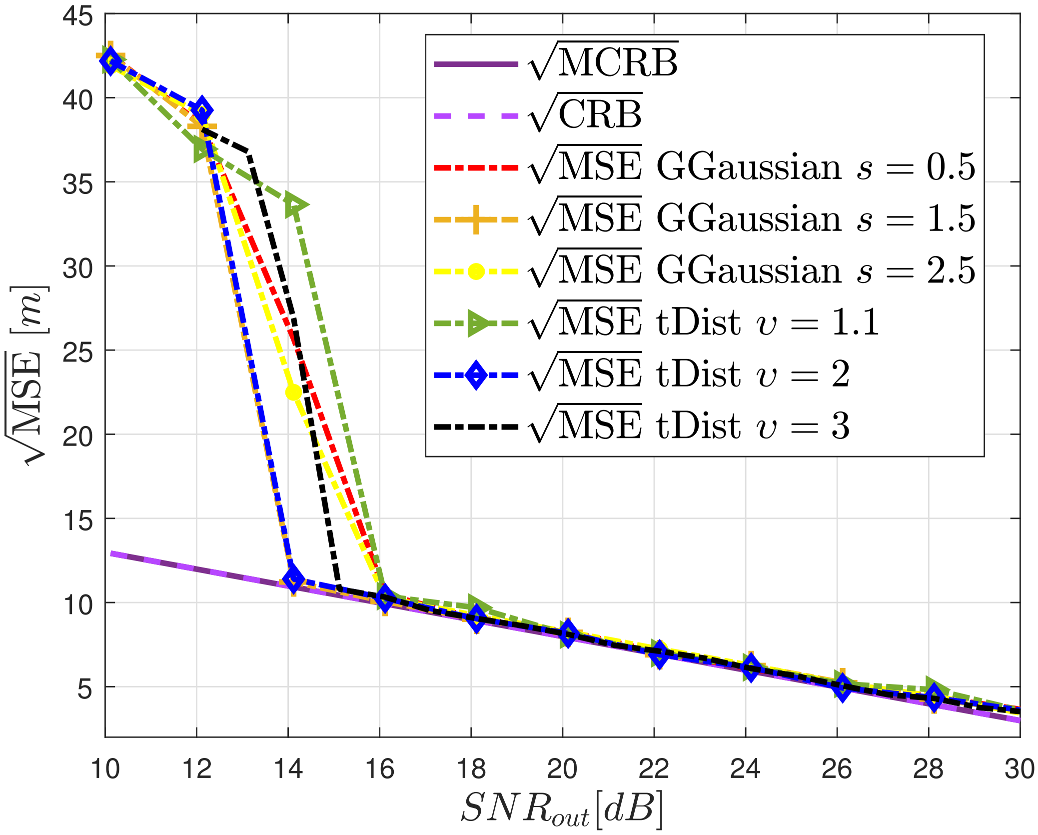

Figure 3: RMSE of the MMLE of the time-delay considering complex centered GG dist. with $s = \{0.5, 1.5, 2.5\}$ and t-dist. with $\upsilon = \{1.1, 2, 3\}$.

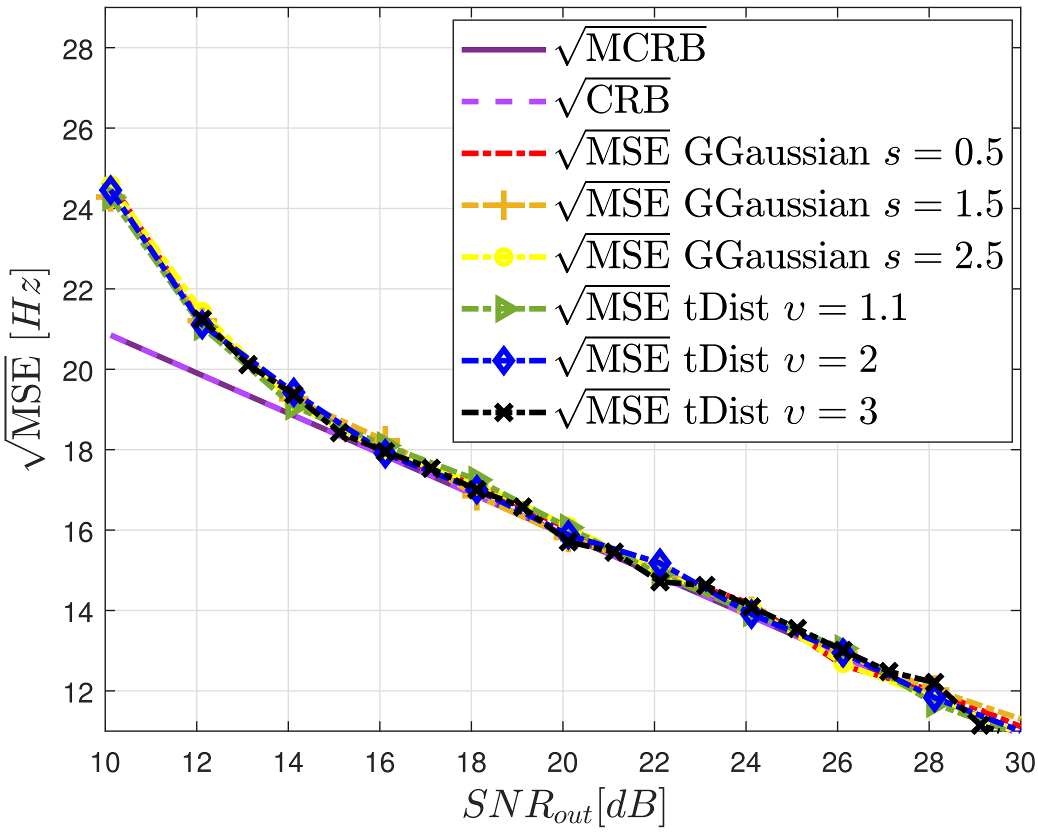

Figure 4: RMSE of the MMLE of the Doppler considering complex centered GG dist. with $s = \{0.5, 1.5, 2.5\}$ and t-dist. with $\upsilon = \{1.1, 2, 3\}$.

View publication stats