The Misspecified and Semiparametric lower bounds and their application to inference problems with Complex Elliptically Symmetric (CES) distributed data

Fulvio Gini and Stefano Fortunati

Dipartimento di Ingegneria dell'Informazione

University of Pisa, Pisa, Italy

Outline – Part I: The Misspecified CRB

1. Motivations and history
2. Formal description of a misspecified problem
3. The Misspecified Cramér-Rao Bound (MCRB)
4. The Mismatched Maximum Likelihood (MML) estimator
5. Simple examples: variance and power estimation
6. Radar applications:
   - Direction of Arrivals (DOAs) estimation
   - Scatter matrix estimation for Complex Elliptically Simmetric (CES) distributed data

Outline – Part II: The Semiparametric CRB

1. Why semiparametric models?
2. CRB for parametric models with finite-dimensional nuisance parameters:
   - ❑ Classical approach,
   - ❑ "Hilbert-space-based" approach.
3. Extension to semiparametric models.
4. Semiparametric interpretation of Real and Complex ES Distributions.
5. Examples.

Why study lower bounds (LBs)?

- **Lower Bounds (LBs)** provide a benchmark of comparison for the performance of any **unbiased** estimator.
- If the performance of a certain estimator achieves a relevant LB, then it is established that no other *unbiased* estimators can do better.
- A LB is said to be **tight** if it reasonably predicts the performance of the Maximum Likelihood (ML) estimator.
- The **Cramér-Rao Bound (CRB)** is an *asymptotically tight* bound for any *unbiased* estimator.

Model misspecification (1/2)

- **Classical "matched" assumption**: the true data model and the model assumed to derive the estimation algorithm are the same, i.e. <u>the model is *correctly specified*</u>.

- All the results on the ML estimator and the CRB rely on this implicit assumption.

- However, much evidence from engineering practice shows that this assumption is often violated.

- **Model misspecification**: the assumed data model (i.e. the data pdf) differs from the true model.

Model misspecification (2/2)

- There are two main reasons for model misspecification:
  1. An **imperfect knowledge** of the true data model that leads to a wrong specification of the data pdf.
  2. The true data model is known but it is **too involved** to pursue the optimal "matched" estimator.
- One may be forced (scenario 1) or may prefer (scenario 2) to derive an estimator by assuming a **simpler** but **misspecified** data model.
- This suboptimal procedure may lead to some degradation in the overall system performance.

Misspecified Lower Bounds (MLBs)

- Assessing the impact of **model misspecification** on the estimation performance is crucial to guarantee the reliability of the (mismatched) estimator.
- **Misspecified LBs** allow the assumed and true models to differ, yet establishing performance limits on estimation error covariance.
- **Misspecified LBs** indicate how the model mis-specification affects estimation performance.

Some history and recent applications (1/3)

- Properties of the **Mismatched ML estimator**: Huber [1] (1967), Akaike [2] (1972) and White [3] (1982).

- Generalization to the **Bayesian** framework: Berk [5] (1966), Bunke and Milhaud [6] (1998), Richmond and Basu [7] (2016).

- **Cramér-Rao inequality** under **model misspecification**: Vuong [4] (1986), Richmond and Horowitz [8] (2015), S. Fortunati, F. Gini, M. S. Greco [11] (2016).

- A tutorial introduction to the **Misspecified CRB** has been proposed in:

S. Fortunati, F. Gini, M. S. Greco and C. D. Richmond, "Performance Bounds for Parameter Estimation under Misspecified Models: Fundamental Findings and Applications," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 142-157, Nov. 2017.

Some history and recent applications (2/3)

- Recent applications of **misspecified LBs**:

1. Direction of Arrivals estimation in sensor arrays [8] and MIMO radars [9].
2. Covariance matrix estimation in non-Gaussian data ([10], [11], [12] and [23]).
3. Radar-communication systems coexistence [7].
4. Waveform parameter estimation in the presence of uncertainty in the propagation model [13].
5. Time of Arrivals (ToA) estimation problem for UWB signals in the presence of interferences [14].

Some history and recent applications (3/3)

6. Sparse Bayesian estimation [27].

7. Spectral estimation [28].

8. Estimation of hybrid sinusoidal FM-PPS signals [29].

Description of a misspecified problem (1/3)

- Let $\mathbf{x}_m \in \mathbb{C}^N$ be an N-dimensional measurement vector.

- Let $p_X(\mathbf{x}_m) \in \mathcal{P}$ its true probability density function (pdf) belonging to a possibly non-parametric model $\mathcal{P}$.

- To characterize the statistical behavior of $\mathbf{x}_m$, we adopt a parametric pdf, say $f_X(\mathbf{x}_m | \boldsymbol{\theta})$ with $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d$.

- The assumed pdf $f_X(\mathbf{x}_m | \boldsymbol{\theta})$ is implicitly assumed to belong to a *parametric* model:

$$\mathcal{F} = \left\{ f_X \middle| f_X(\mathbf{x}_m | \boldsymbol{\theta}), \boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d \right\}$$

Description of a misspecified problem (2/3)

- The classical "matched" estimation theory requires the existence of a parameter vector $\bar{\boldsymbol{\theta}} \in \Theta$ such that:

$$p_X(\mathbf{x}_m) = f_X(\mathbf{x}_m | \bar{\boldsymbol{\theta}}) \text{ or, equivalently, } p_X(\mathbf{x}_m) \in \mathcal{F}$$

- If this assumption is violated, the model is **misspecified**.

Model misspecification:

$f_X(\mathbf{x}_m | \boldsymbol{\theta})$ differs from $p_X(\mathbf{x}_m)$ for every $\boldsymbol{\theta} \in \Theta$.

$$p_X(\mathbf{x}_m) \neq f_X(\mathbf{x}_m | \boldsymbol{\theta}), \forall \boldsymbol{\theta} \in \Theta \quad \longrightarrow \quad \mathcal{P} \not\subset \mathcal{F}$$

Description of a misspecified problem (3/3)

- Suppose to collect M *independent, identically distributed* (i.i.d.) N-dimensional measurement vectors:

$$\mathbf{x} = \{\mathbf{x}_m\}_{m=1}^{M}, \quad \mathbf{x}_m \sim p_X(\mathbf{x}_m), \; m = 1,\dots,M$$

- Due to the independence assumption, the true joint pdf of the dataset **x** is:

$$p_X(\mathbf{x}) = \prod_{m=1}^{M} p_X(\mathbf{x}_m)$$

- The assumed joint pdf of the dataset **x** is:

$$f_X(\mathbf{x}|\boldsymbol{\theta}) = \prod_{m=1}^{M} f_X(\mathbf{x}_m|\boldsymbol{\theta})$$

Two fundamental questions

- This misspecified scenario raises two main questions:

  1. Is it still possible to derive LBs on the error covariance of any mismatched estimator of the parameter vector $\boldsymbol{\theta}$?

  2. How will the classical statistical properties of an estimator, e.g. *unbiasedness*, *consistency* and *efficiency*, change in this misspecified model framework?

- The **Misspecified Cramér-Rao Bound (MCRB)** provides answers to these questions.

Regular models

- As for the classical CRB, in order to guarantee the existence of the MCRB, some regularity conditions need to be imposed.

- Specifically, the assumed parametric model $\mathcal{F}$ has to be **regular** with respect to (wrt) the true one $\mathcal{P}$.

- Among the rather technical assumptions that $\mathcal{F}$ has to satisfy to be regular wrt $\mathcal{P}$, the most important is:

  - Existence and uniqueness of the pseudo-true parameter vector $\boldsymbol{\theta}_0$

The pseudo-true parameter vector

- If $\mathcal{F}$ is regular wrt $\mathcal{P}$, then there exist a unique interior point $\boldsymbol{\theta}_0$ of Θ, such that:

$$\boldsymbol{\theta}_0 \triangleq \arg\min_{\boldsymbol{\theta}\in\Theta}\left\{-E_p\left\{\ln f_X\left(\mathbf{x}_m\,|\,\boldsymbol{\theta}\right)\right\}\right\} = \arg\min_{\boldsymbol{\theta}\in\Theta}\left\{D\left(p_X \,\|\, f_X\right)\right\}$$

where $E_p\{\cdot\}$ is the expectation operator wrt the true pdf $p_X(\mathbf{x}_m)$ and

$$D\left(p_X \,\|\, f_X\right) \triangleq \int \ln\left(\frac{p_X(\mathbf{x}_m)}{f_X(\mathbf{x}_m\,|\,\boldsymbol{\theta})}\right) p_X(\mathbf{x}_m)\,d\mathbf{x}_m$$

is the **Kullback-Leibler divergence (KLD)** between the true pdf and the assumed pdf.

Generalization of the FIM

- The vector $\boldsymbol{\theta}_0$ is the point the minimizes the KLD between the true pdf $p_X(\mathbf{x}_m)$ and the assumed pdf $f_X(\mathbf{x}_m|\boldsymbol{\theta})$.

- Let $\mathbf{A}_{\boldsymbol{\theta}_0}$ be the matrix whose entries are given by:

$$\left[\mathbf{A}_{\boldsymbol{\theta}_0}\right]_{ij} \triangleq E_p\left\{ \frac{\partial^2}{\partial\theta_i\partial\theta_j} \ln f_X\left(\mathbf{x}_m|\boldsymbol{\theta}\right)\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \right\}$$

If $\mathcal{F}$ is regular wrt $\mathcal{P}$, then the matrix $\mathbf{A}_{\boldsymbol{\theta}_0}$ is non singular

- In the matrix $\mathbf{A}_{\boldsymbol{\theta}_0}$ we can recognize a sort of generalization of the Fisher Information Matrix (FIM).

Another generalization of the FIM

- The second generalization of the FIM is given by:

$$\left[\mathbf{B}_{\boldsymbol{\theta}_0}\right]_{ij} \triangleq E_p\left\{ \left.\frac{\partial \ln f_X\left(\mathbf{x}_m|\boldsymbol{\theta}\right)}{\partial \theta_i}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \cdot \left.\frac{\partial \ln f_X\left(\mathbf{x}_m|\boldsymbol{\theta}\right)}{\partial \theta_j}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \right\}.$$

- If the model is **misspecified**, i.e. it does not exists a parameter vector $\overline{\boldsymbol{\theta}} \in \Theta$ such that $p_X(\mathbf{x}_m) = f_X(\mathbf{x}_m|\overline{\boldsymbol{\theta}})$, then: $\mathbf{B}_{\boldsymbol{\theta}_0} \neq -\mathbf{A}_{\boldsymbol{\theta}_0}$

The MS-unbiasedness property

- The Misspecified (MS)-unbiasedness property generalizes the classical notion of unbiased estimators.

- In the misspecified model framework, unbiasedness is defined wrt the pseudo-true parameter vector $\boldsymbol{\theta}_0$.

- Let $\mathbf{x} = \{\mathbf{x}_m\}_{m=1}^{M}$ be the available dataset and let $\hat{\boldsymbol{\theta}}(\mathbf{x})$ be an estimator derived by assuming the misspecified model $\mathcal{F}$.

$\hat{\boldsymbol{\theta}}(\mathbf{x})$ is an **MS-unbiased** estimator iff:

$$E_p\left\{\hat{\boldsymbol{\theta}}(\mathbf{x})\right\} = \int \hat{\boldsymbol{\theta}}(\mathbf{x}) p_X(\mathbf{x}) d\mathbf{x} = \boldsymbol{\theta}_0$$

The Misspecified Cramér-Rao Bound (1/2)

- Let $\mathcal{F}$ be a misspecified parametric model that is regular wrt a true model $\mathcal{P}$.

Theorem 1 (Vuong 1986): Let $\hat{\boldsymbol{\theta}}(\mathbf{x})$ be an MS-unbiased estimator derived under the misspecified model $\mathcal{F}$ from a dataset $\mathbf{x}=\{\mathbf{x}_m\}_{m=1}^{M}$. Then, for every possible $p_X(\mathbf{x}_m)\in\mathcal{P}$:

$$\mathbf{C}_p\left(\hat{\boldsymbol{\theta}}(\mathbf{x}),\boldsymbol{\theta}_0\right)\geq\frac{1}{M}\mathbf{A}_{\boldsymbol{\theta}_0}^{-1}\mathbf{B}_{\boldsymbol{\theta}_0}\mathbf{A}_{\boldsymbol{\theta}_0}^{-1}\triangleq\mathrm{MCRB}\left(\boldsymbol{\theta}_0\right)$$

where

$$\mathbf{C}_p\left(\hat{\boldsymbol{\theta}}(\mathbf{x}),\boldsymbol{\theta}_0\right)\triangleq E_p\left\{\left(\hat{\boldsymbol{\theta}}(\mathbf{x})-\boldsymbol{\theta}_0\right)\left(\hat{\boldsymbol{\theta}}(\mathbf{x})-\boldsymbol{\theta}_0\right)^T\right\}$$

is the error covariance matrix of the mismatched estimator.

[4] Q. H. Vuong, “Cramér-Rao bounds for misspecified models”, *Working paper 652, Division of the Humanities and Social Sciences* , Caltech, October 1986.

The Misspecified Cramér-Rao Bound (2/2)

- The MCRB is a local lower bound (LB) on the error variance of any MS-unbiased estimator of the pseudo-true parameter vector $\boldsymbol{\theta}_0$.

- MCRB for constrained estimation problem [20]:

  S. Fortunati, F. Gini, M. S. Greco, "The Constrained Misspecified Cramér-Rao Bound," *IEEE Signal Process. Letters*, Vol. 23, No. 5, pp. 718-721, May 2016.

- Extension to complex parameters [21]:

  S. Fortunati, "Misspecified Cramér-Rao Bounds for Complex Unconstrained and Constrained Parameters," *EUSIPCO 2017*, Kos, Greece, 28 Aug. 2017–2 Sept. 2017.

MCRB vs CRB

- If the model is **correctly specified**, i.e. if there exists a parameter vector $\bar{\boldsymbol{\theta}} \in \Theta$, such that $p_X(\mathbf{x}_m) = f_X(\mathbf{x}_m | \bar{\boldsymbol{\theta}})$, then:

$$\boldsymbol{\theta}_0 = \bar{\boldsymbol{\theta}} \text{ and } \mathbf{B}_{\boldsymbol{\theta}_0} = -\mathbf{A}_{\boldsymbol{\theta}_0} = -\mathbf{A}_{\bar{\boldsymbol{\theta}}} \quad \Rightarrow \quad \mathrm{MCRB} = \mathrm{CRB}$$

$$E_p\left\{\left(\hat{\boldsymbol{\theta}}(\mathbf{x}) - \bar{\boldsymbol{\theta}}\right)\left(\hat{\boldsymbol{\theta}}(\mathbf{x}) - \bar{\boldsymbol{\theta}}\right)^T\right\} \geq -\frac{1}{M}\mathbf{A}_{\bar{\boldsymbol{\theta}}}^{-1} = \mathrm{MCRB}\left(\bar{\boldsymbol{\theta}}\right) = \mathrm{CRB}\left(\bar{\boldsymbol{\theta}}\right)$$

The misspecified framework is consistent with the classical matched theory!

The Mismatched ML estimator (MML) (1/3)

- As before, suppose to collect M *independent, identically distributed* (i.i.d.) N-dimensional measurement vectors:

$$\mathbf{x} = \{\mathbf{x}_m\}_{m=1}^{M}, \quad \mathbf{x}_m \sim p_X(\mathbf{x}_m), m = 1, \ldots, M$$

- The log-likelihood function for the dataset **x** under a misspecified model $\mathcal{F}$ is given by:

$$L_M(\boldsymbol{\theta}) \triangleq \frac{1}{M} \sum_{m=1}^{M} \ln f_X(\mathbf{x}_m | \boldsymbol{\theta}), \quad \mathbf{x}_m \sim p_X(\mathbf{x}_m)$$

- The **MML estimator** is the point that maximizes $L_M(\boldsymbol{\theta})$:

$$\hat{\boldsymbol{\theta}}_{MML}(\mathbf{x}) \triangleq \underset{\boldsymbol{\theta} \in \Theta}{\arg\max}\, L_M(\boldsymbol{\theta}) = \underset{\boldsymbol{\theta} \in \Theta}{\arg\max} \sum_{m=1}^{M} \ln f_X(\mathbf{x}_m | \boldsymbol{\theta})$$

The Mismatched ML estimator (MML) (2/3)

Theorem 2 (Huber 1967, White 1982): Under suitable regularity conditions, it can be shown that:

$$\hat{\boldsymbol{\theta}}_{MML}(\mathbf{x}) \underset{M\to\infty}{\overset{a.s.}{\rightarrow}} \boldsymbol{\theta}_0.$$

Moreover:

$$\sqrt{M}\left(\hat{\boldsymbol{\theta}}_{MML}(\mathbf{x})-\boldsymbol{\theta}_0\right) \underset{M\to\infty}{\overset{d}{\sim}} \mathcal{N}\left(\mathbf{0},\mathbf{A}_{\boldsymbol{\theta}_0}^{-1}\mathbf{B}_{\boldsymbol{\theta}_0}\mathbf{A}_{\boldsymbol{\theta}_0}^{-1}\right).$$

Huber "sandwich" covariance matrix = MCRB

- $\underset{M\to\infty}{\overset{a.s.}{\rightarrow}}$ indicates the *almost sure* (*a.s.*) convergence.
- $\underset{M\to\infty}{\overset{d}{\sim}}$ indicates the convergence *in distribution*.

The Mismatched ML estimator (MML) (3/3)

- The MML estimator is asymptotically MS-unbiased and its asymptotic error covariance is equal to the MCRB, i.e. it is an *asymptotically efficient* estimator wrt the MCRB.

- Consistent with the *classical* "matched" ML estimator:

If the model $\mathcal{F}$ is correctly specified, i.e. if there exists

$$\bar{\boldsymbol{\theta}} \in \Theta \text{ such that: } p_X(\mathbf{x}_m) = f_X(\mathbf{x}_m | \bar{\boldsymbol{\theta}})$$

then:

$$\hat{\boldsymbol{\theta}}_{ML}(\mathbf{x}) \underset{M \to \infty}{\overset{a.s.}{\rightarrow}} \boldsymbol{\theta}_0 = \bar{\boldsymbol{\theta}}$$

CRB

$$\sqrt{M}\left(\hat{\boldsymbol{\theta}}_{ML}(\mathbf{x}) - \bar{\boldsymbol{\theta}}\right) \underset{M \to \infty}{\overset{d}{\sim}} \mathcal{N}\left(\mathbf{0}, -\mathbf{A}_{\bar{\boldsymbol{\theta}}}^{-1}\right)$$

Consistent estimate of the MCRB

- Let us define the following data-dependent matrices:

$$[\mathbf{A}_M(\boldsymbol{\theta})]_{ij} \triangleq M^{-1}\sum_{m=1}^{M}\frac{\partial^2 \ln f_X(\mathbf{x}_m|\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}$$

$$[\mathbf{B}_M(\boldsymbol{\theta})]_{ij} \triangleq M^{-1}\sum_{m=1}^{M}\frac{\partial \ln f_X(\mathbf{x}_m|\boldsymbol{\theta})}{\partial\theta_i}\cdot\frac{\partial \ln f_X(\mathbf{x}_m|\boldsymbol{\theta})}{\partial\theta_j}$$

$$\mathbf{C}_M(\boldsymbol{\theta}) \triangleq [\mathbf{A}_M(\boldsymbol{\theta})]^{-1}\mathbf{B}_M(\boldsymbol{\theta})[\mathbf{A}_M(\boldsymbol{\theta})]^{-1}$$

- It can be shown (see [3, Theo 3.2]) that:

$$\left[\mathbf{C}_M(\hat{\boldsymbol{\theta}}_{MML})\right]_{i,j} \underset{M\to\infty}{\overset{a.s.}{\to}} \left[\mathbf{C}_{\boldsymbol{\theta}_0}\right]_{i,j} = \left[\mathrm{MCRB}(\boldsymbol{\theta}_0)\right]_{i,j} \quad \forall i,j = 1,\ldots,\dim(\boldsymbol{\theta}_0)$$

A test for model misspecification

- It is possible to infer from the collected dataset whether or not the assumed model $\mathcal{F}$ is correctly specified.
- Recall that, if $\mathcal{F}$ is correctly specified, then $\mathbf{B}_{\boldsymbol{\theta}} = -\mathbf{A}_{\boldsymbol{\theta}}$.
- Since $\mathbf{A}_{\boldsymbol{\theta}}$ and $\mathbf{B}_{\boldsymbol{\theta}}$ are unobservable, we can exploit their sample estimate to implement a composite hypothesis testing:

$$\begin{cases} H_0: & \mathbf{A}_M(\hat{\boldsymbol{\theta}}_{MML}) + \mathbf{B}_M(\hat{\boldsymbol{\theta}}_{MML}) = \mathbf{0} \\ H_1: & \mathbf{A}_M(\hat{\boldsymbol{\theta}}_{MML}) + \mathbf{B}_M(\hat{\boldsymbol{\theta}}_{MML}) \neq \mathbf{0} \end{cases}$$

- A Wald-type test can be derived to discriminate between the two hypotheses: *correct specification* (H_0) *vs model misspecification* (H_1) (see [3, Sec. 4]).

Example 1: Variance estimation (1/6)

- **Problem**: we want to estimate the variance of a Gaussian data set in the presence of misspecified mean value, e.g. we erroneously assume that the mean value is zero.

- True data set:

$$\mathbf{x} = \{x_m\}_{m=1}^{M}, \quad x_m \sim p_X(x_m) \equiv \mathcal{N}(\bar{\mu}, \bar{\sigma}^2), \quad \bar{\mu} \neq 0$$

- Assumed data model:

$$\mathcal{F} = \left\{ f_X \middle| f_X(x_m|\theta) \equiv \mathcal{N}(0,\theta) \; \forall \theta \in \mathbb{R}^+ \right\}$$

- Note that, as long as $\bar{\mu} \neq 0$,

$$p_X(x_m) \notin \mathcal{F}$$

Example 1: Variance estimation (2/6)

- Is the misspecified model $\mathcal{F}$ regular wrt $p(x_m)$?

- We have to check if:

1. there exists the pseudo-true parameter θ_0;
2. the matrix $\mathbf{A}_{\boldsymbol{\theta}_0}$ is non singular.

- The KLD can be expressed as:

$$D\left(p_X \| f_X\right) = \frac{\bar{\mu}^2}{2\theta} + \frac{1}{2}\left(\frac{\bar{\sigma}^2}{\theta} - 1 - \ln\frac{\bar{\sigma}^2}{\theta}\right)$$

- Its minimum point, i.e. the pseudo-true parameter, *exists* and is *unique*:

$$\theta_0 = \bar{\sigma}^2 + \bar{\mu}^2, \quad \bar{\mu} \neq 0$$

Example 1: Variance estimation (3/6)

- The scalar A_{θ_0} can be evaluated as:

$$A_{\theta_0} \triangleq E_p\left\{ \left.\frac{\partial^2 \ln f_X(x_m \mid \theta)}{\partial \theta^2}\right|_{\theta=\theta_0} \right\} = \frac{1}{2\theta_0^2} - \frac{1}{\theta_0^3} E_p\left\{x_m^2\right\} = -\frac{1}{2\theta_0^2}$$

that is always different from zero, since $\bar{\mu} \neq 0, \bar{\sigma}^2 \in \mathbb{R}^+$.

Since θ_0 exists and is unique and the scalar $A_{\theta_0} \neq 0$, then $\mathcal{F}$ is regular wrt $p\left(x_m\right)$.

Example 1: Variance estimation (4/6)

- The scalar B_{θ_0} can be evaluated as:

$$B_{\theta_0} \triangleq E_p\left\{ \left(\frac{\partial \ln f_X(x_m|\theta)}{\partial \theta} \right)^2 \Bigg|_{\theta=\theta_0} \right\} = \frac{2\bar{\sigma}^4 + 4\bar{\sigma}^2\bar{\mu}^2}{4\theta_0^4}$$

- By collecting the previous results, the MCRB is given by:

$$\mathrm{MCRB}(\theta_0) = \frac{2\bar{\sigma}^4}{M} + \frac{4\bar{\sigma}^2\bar{\mu}^2}{M} \geq \mathrm{CRB}(\bar{\sigma}^2) = \frac{2\bar{\sigma}^4}{M}$$

- The MCRB is always greater than the CRB, as expected.

- The MCRB equates the CRB when there is no model misspecification, i.e. when $\bar{\mu} = 0$.

Example 1: Variance estimation (5/6)

- Regarding the MML estimator, we have that:

$$\hat{\theta}_{MML}(\mathbf{x}) = \frac{1}{M}\sum_{m=1}^{M} x_m^2 \underset{M\to\infty}{\overset{a.s.}{\longrightarrow}} \theta_0 = \bar{\sigma}^2 + \bar{\mu}^2 \neq \bar{\sigma}^2$$

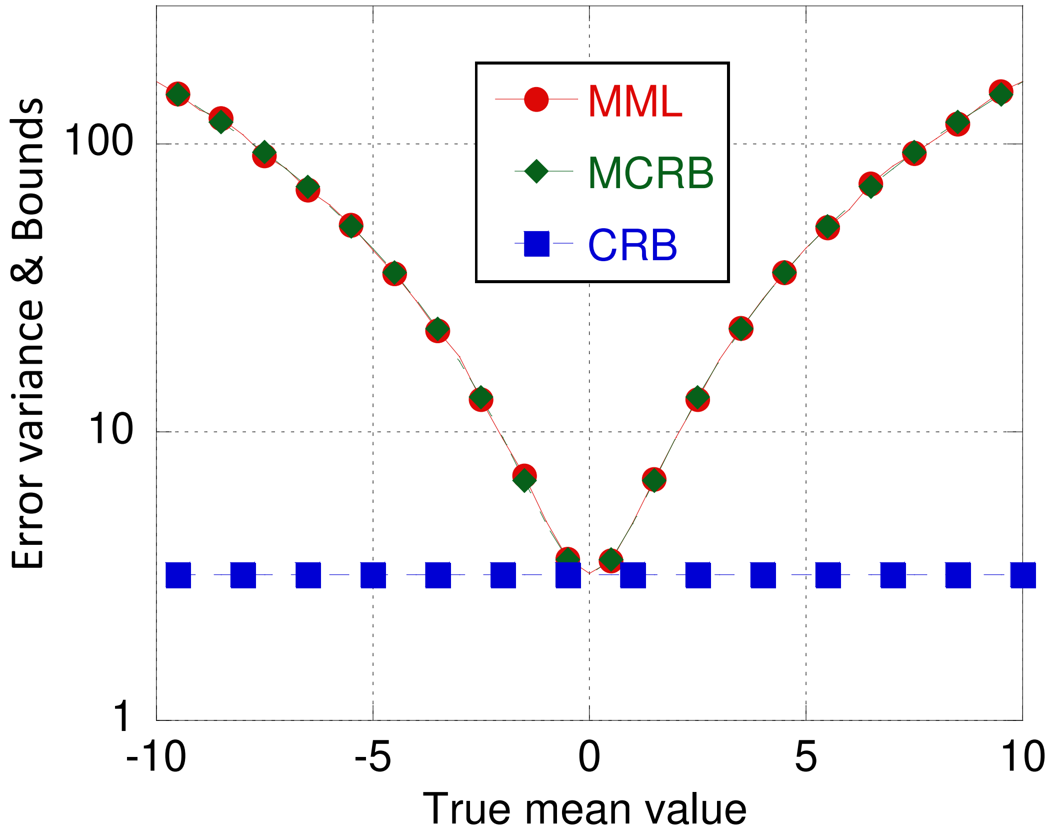
- The MML estimator is not consistent, since it converges to a value different from the true variance.

- However, the MML is MS-unbiased, since:

$$E_p\left\{\hat{\theta}_{MML}(\mathbf{x})\right\} = E_p\left\{\frac{1}{M}\sum_{m=1}^{M} x_m^2\right\} = \bar{\sigma}^2 + \bar{\mu}^2 = \theta_0$$

- Hence, the error variance of this MML estimator is lower bounded by the MCRB.

Example 1: Variance estimation (6/6)



$\bar{\sigma}^2 = 4$

$M = 10$

- MCRB = CRB when $\bar{\mu} = 0$: matched case.
- The MML estimator is efficient wrt the MCRB.

Example 2: Power estimation (1/6)

- **Problem**: we want to estimate the statistical power of a multivariate Gaussian data set under misspecification of the correlation structure.

- True dataset:

$$\mathbf{x} = \{\mathbf{x}_m\}_{m=1}^{M} \qquad p_X(\mathbf{x}_m) \equiv \mathcal{N}\left(\mathbf{0}, \bar{\sigma}^2 \mathbf{\Sigma}\right) \in \mathcal{P} \qquad [\mathbf{\Sigma}]_{ij} = \rho^{|i-j|}$$

- Assumed data model:

$$\mathcal{F} = \left\{ f_X \mid f_X(\mathbf{x}_m \mid \theta) \equiv \mathcal{N}\left(\mathbf{0}, \theta \mathbf{I}_N\right) \ \forall \theta \in \mathbb{R}^+ \right\}$$

- Note that, as long as $\rho \neq 0$,

$$p_X\left(\mathbf{x}_m\right) \notin \mathcal{F}$$

Example 2: Power estimation (2/6)

- Is the misspecified model $\mathcal{F}$ regular wrt $p(\mathbf{x}_m)$?

- As before, we have to check if:

1. there exists the pseudo-true parameter θ_0;
2. the scalar $A_{\boldsymbol{\theta}_0}$ is non singular.

- The KLD can be expressed as:

$$D\left(p_X \| f_X\right) = \frac{1}{2}\left[\operatorname{tr}(\theta^{-1}\bar{\sigma}^2\mathbf{\Sigma}) - N + \ln\theta - \ln\det(\bar{\sigma}^2\mathbf{\Sigma})\right]$$

- Its minimum point, i.e. the pseudo-true parameter, *exists* and is *unique*: $\theta_0 = \bar{\sigma}^2$

Example 2: Power estimation (3/6)

- The scalar A_{θ_0} can be evaluated as:

$$A_{\theta_0} \triangleq E_p\left\{ \frac{\partial^2 \ln f_X(\mathbf{x}_m|\theta)}{\partial \theta^2}\bigg|_{\theta=\theta_0} \right\} = \frac{N}{2\theta_0^2} - \frac{1}{\theta_0^3} E_p\left\{\mathbf{x}_m^T \mathbf{x}_m\right\} = -\frac{N}{2\bar{\sigma}^4}$$

that is always different from zero, since $\bar{\sigma}^2 \in \mathbb{R}^+$.

Since θ_0 exists and is unique and the scalar $A_{\theta_0} \neq 0$, then $\mathcal{F}$ is regular wrt $p(\mathbf{x}_m)$.

Example 2: Power estimation (4/6)

- The scalar B_{θ_0} can be evaluated as:

$$B_{\theta_0} = \frac{N\theta_0^2 + E_p\left\{(\mathbf{x}_m^T\mathbf{x}_m)^2\right\} - 2N\theta_0 E_p\left\{\mathbf{x}_m^T\mathbf{x}_m\right\}}{4\theta_0^4} = \frac{\mathrm{tr}\left(\boldsymbol{\Sigma}^2\right)}{2\bar{\sigma}^4}$$

- By collecting the previous results, the MCRB is given by:

$$\mathrm{MCRB}(\theta_0) = \mathrm{MCRB}(\bar{\sigma}^2) = \frac{2\bar{\sigma}^4}{MN^2}\mathrm{tr}\left(\boldsymbol{\Sigma}^2\right) \geq \mathrm{CRB}(\bar{\sigma}^2) = \frac{2\bar{\sigma}^4}{MN}$$

- The MCRB is always greater than the CRB, as expected.

- MCRB = CRB when there is no model misspecification, i.e. when $\rho = 0$.

Example 2: Power estimation (5/6)

- Regarding the MML estimator, we have that:

$$\hat{\theta}_{MML}(\mathbf{x}) = \frac{1}{M}\sum_{m=1}^{M}\frac{\mathbf{x}_m^T\mathbf{x}_m}{N} \xrightarrow[M\to\infty]{a.s.} \theta_0 = \frac{E_p\{\mathbf{x}_m^T\mathbf{x}_m\}}{N} = \frac{\bar{\sigma}^2}{N}\mathrm{tr}(\mathbf{\Sigma}) = \bar{\sigma}^2$$

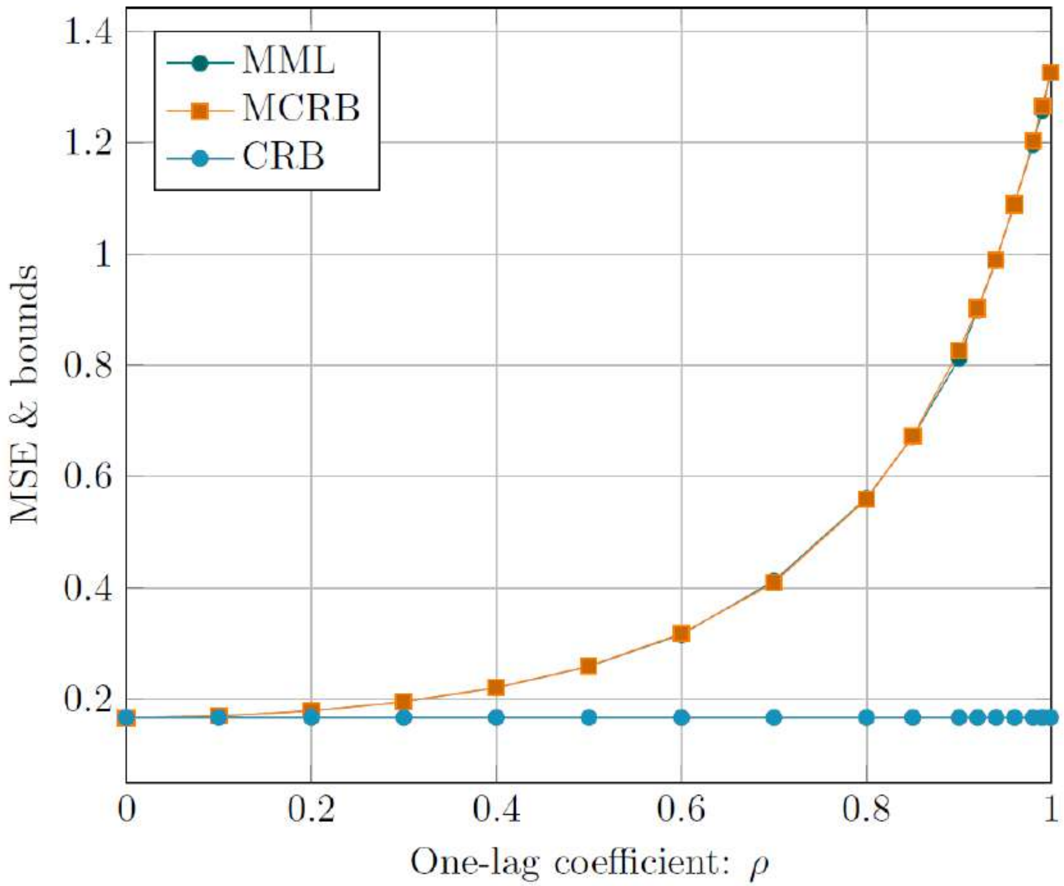
- The MML estimator is consistent, since it converges to the true value of the statistical power.

- Moreover, the MML is MS-unbiased, since:

$$E_p\{\hat{\theta}_{MML}(\mathbf{x})\} = \frac{1}{M}\sum_{m=1}^{M}\frac{E_p\{\mathbf{x}_m^T\mathbf{x}_m\}}{N} = \bar{\sigma}^2 = \theta_0$$

- Hence, the error variance (and the MSE) of this MML estimator is lower bounded by the MCRB.

Example 2: Power estimation (6/6)



$\bar{\sigma}^2 = 4$

$M = 3N$

$N = 8$

- MCRB = CRB when $\rho = 0$: matched case.
- The MML estimator is efficient wrt the MCRB.

Applications: DOA estimation (1/8)

- **Problem**: we want to estimate the Directions of Arrival (DOAs) of plane-waves signals by means of an array of sensors.
- This is a core research within the SP community [15].
- The fundamental prerequisite for any DOA estimation algorithm is that the positions of the sensors in the array are known exactly.
- Many authors have investigated the impact on the DOA estimation performance of an imperfect knowledge of the sensor positions or of the misscalibration of the array itself (see e.g. [16] and [17]).

Applications: DOA estimation (2/8)

- Some authors have proposed *hybrid* or *modified* CRB with the aim to predict the lowest MSE of the DOA estimators in the presence of the position uncertainties ([18], [19]).
- All these classical approaches, although reasonable, are application-dependent and not general.
- The application of the general misspecified estimation framework to DOA estimation problems has been firstly proposed by Richmond and Horowitz in their seminal paper:

[8] Richmond, C.D.; Horowitz, L.L., "Parameter Bounds on Estimation Accuracy Under Model Misspecification," *IEEE Trans. on Signal Process.*, vol.63, no.9, pp.2263-2278, 2015

Applications: DOA estimation (3/8)

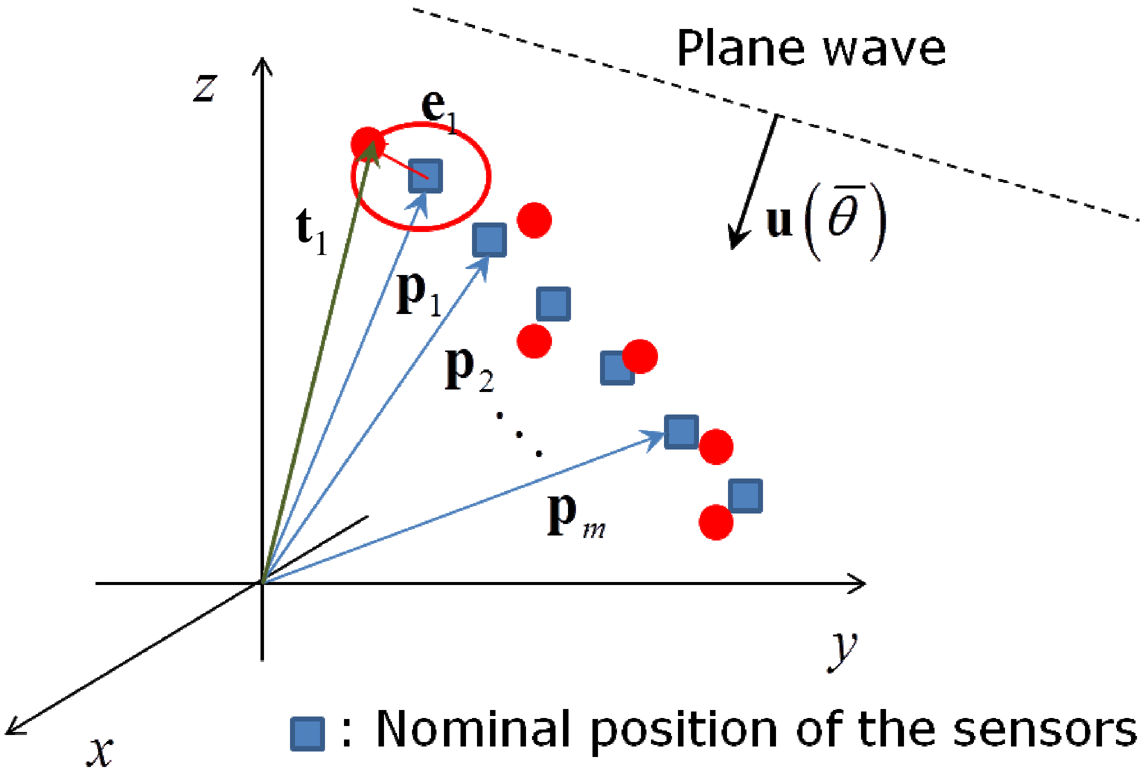
- Following [8], consider a ULA of M sensors and a single plane wave signal impinging on the array from a conic angle $\bar{\theta}$.

- Due to array misscalibration, the true position vector of the m^{th} sensor is known up to an error term modeled as a zero-mean Gaussian random vector $\mathbf{e}_m \sim \mathcal{N}(0, \sigma_e^2 \mathbf{I}_3)$.

- Define as $\mathbf{u}(\bar{\theta})$ the unit vector pointing at the direction of the impinging plane wave.

- Define $\mathbf{k}_{\bar{\theta}} \triangleq (2\pi/\lambda)\mathbf{u}(\bar{\theta})$ where λ is the wavelength.

Applications: DOA estimation (4/8)



■ : Nominal position of the sensors $\mathbf{p}_m$

● : True position of the sensors $\mathbf{t}_m = \mathbf{p}_m + \mathbf{e}_m$

- The true (perturbed) steering vector is given by:

$$[\mathbf{d}(\bar{\theta})]_m = \exp\left(j\mathbf{k}_{\bar{\theta}}^T (\mathbf{p}_m + \mathbf{e}_m) \right), \quad m = 1, \ldots, M$$

Applications: DOA estimation (5/8)

- The signal received at the *m*th sensor is:

$$x_m = \bar{s} \cdot [\mathbf{d}(\bar{\theta})]_m + [\mathbf{c}]_m, \quad m = 1, \ldots, M$$

- $\bar{s}$ is the deterministic unknown complex amplitude.

- $\mathbf{c} = \mathbf{n} + \mathbf{j}$ is the disturbance vector composed of a white Gaussian noise $\mathbf{n}$ and possibly also of a jammer $\mathbf{j}$.

- Given particular realizations of $\mathbf{e}_m$, the disturbance can be modeled as:

$$\mathbf{c} \sim \mathcal{N}\left(0, \sigma_n^2 \mathbf{I}_M + \sigma_j^2 \mathbf{d}(\theta_j)\mathbf{d}^H(\theta_j)\right)$$

Power and DOA of the jammer

Applications: DOA estimation (6/8)

- Since the particular realizations of $\mathbf{e}_m$ is generally unknown, one may decide to assume the <u>nominal steering vector</u> in the estimation algorithm:

$$[\mathbf{v}(\theta)]_m = \exp(j\mathbf{k}_\theta^T \mathbf{p}_m), \quad m = 1, \ldots, M$$

- The true (but unknown) data model is:

$$p_X(\mathbf{x}) = \mathcal{N}\left(\bar{s}\mathbf{d}(\bar{\theta}), \sigma_n^2 \mathbf{I}_M + \sigma_j^2 \mathbf{d}(\theta_j)\mathbf{d}^H(\theta_j)\right) \in \mathcal{P}$$

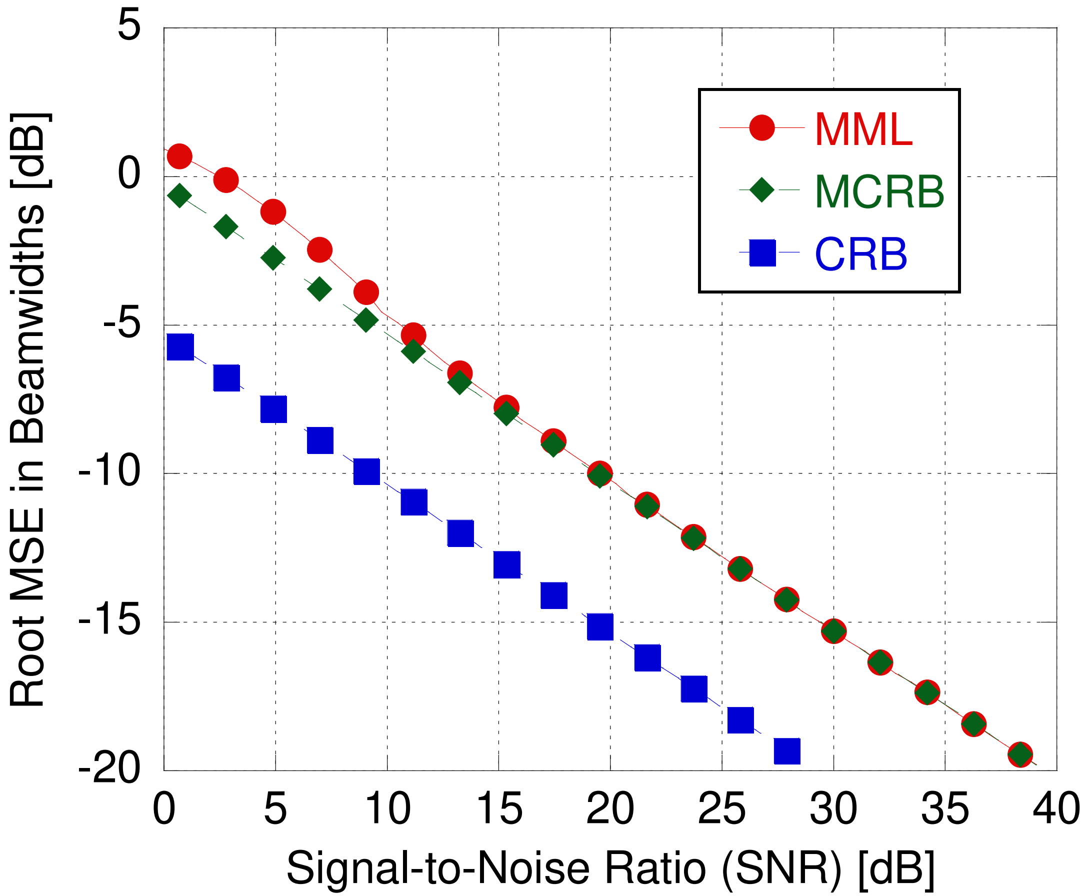
- The assumed parametric data model is:

$$\mathcal{F} = \left\{ f_X \,\middle|\, f_X(\mathbf{x} \mid s, \theta) \equiv \mathcal{N}\left(s\mathbf{v}(\theta), \sigma_n^2 \mathbf{I}_M + \sigma_j^2 \mathbf{v}(\theta_j)\mathbf{v}^H(\theta_j)\right) \right\}$$

where $s \in \mathbb{C}, \theta \in [0, 2\pi)$

Applications: DOA estimation (7/8)

- The MCRB can predict how large is the performance loss in the estimation of $\bar{\theta}$ due to this model mismatch.



Applications: DOA estimation (8/8)

- The MCRB accurately predicts performance of the Mismatched ML (MML) estimator.

- If the system goal is a 10-to-1 beamsplit ratio, i.e. -10dB RMSE in beamwidths, then this could be accomplished with an *SNR* of ~10dB when the model is perfectly known.

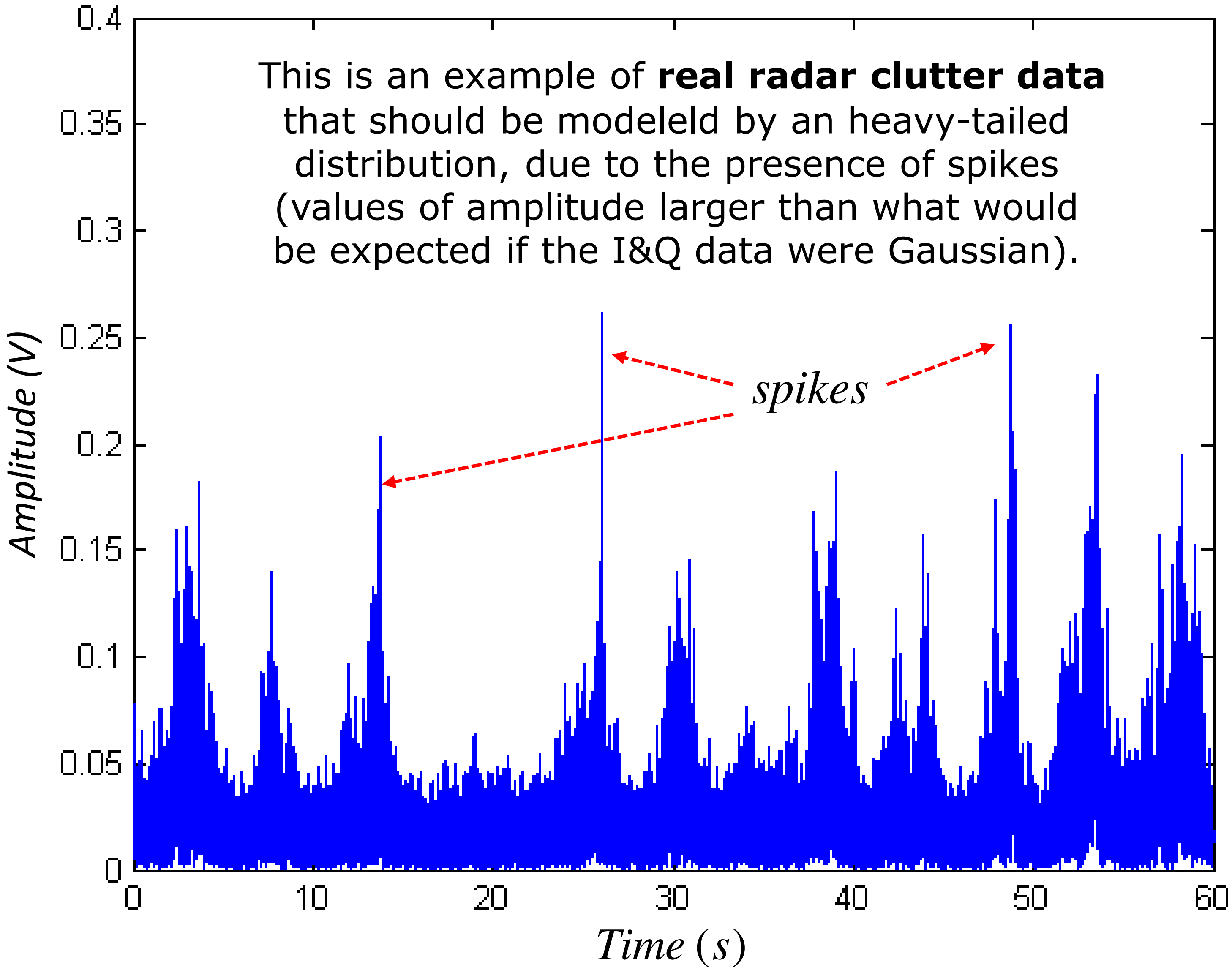
- However, not knowing precisely the true sensor positions requires an additional ~10dB of *SNR* to achieve the same goal [8].

[8] Richmond, C.D.; Horowitz, L.L., "Parameter Bounds on Estimation Accuracy Under Model Misspecification," *IEEE Trans. on Signal Process.*, vol.63, no.9, pp.2263-2278, 2015

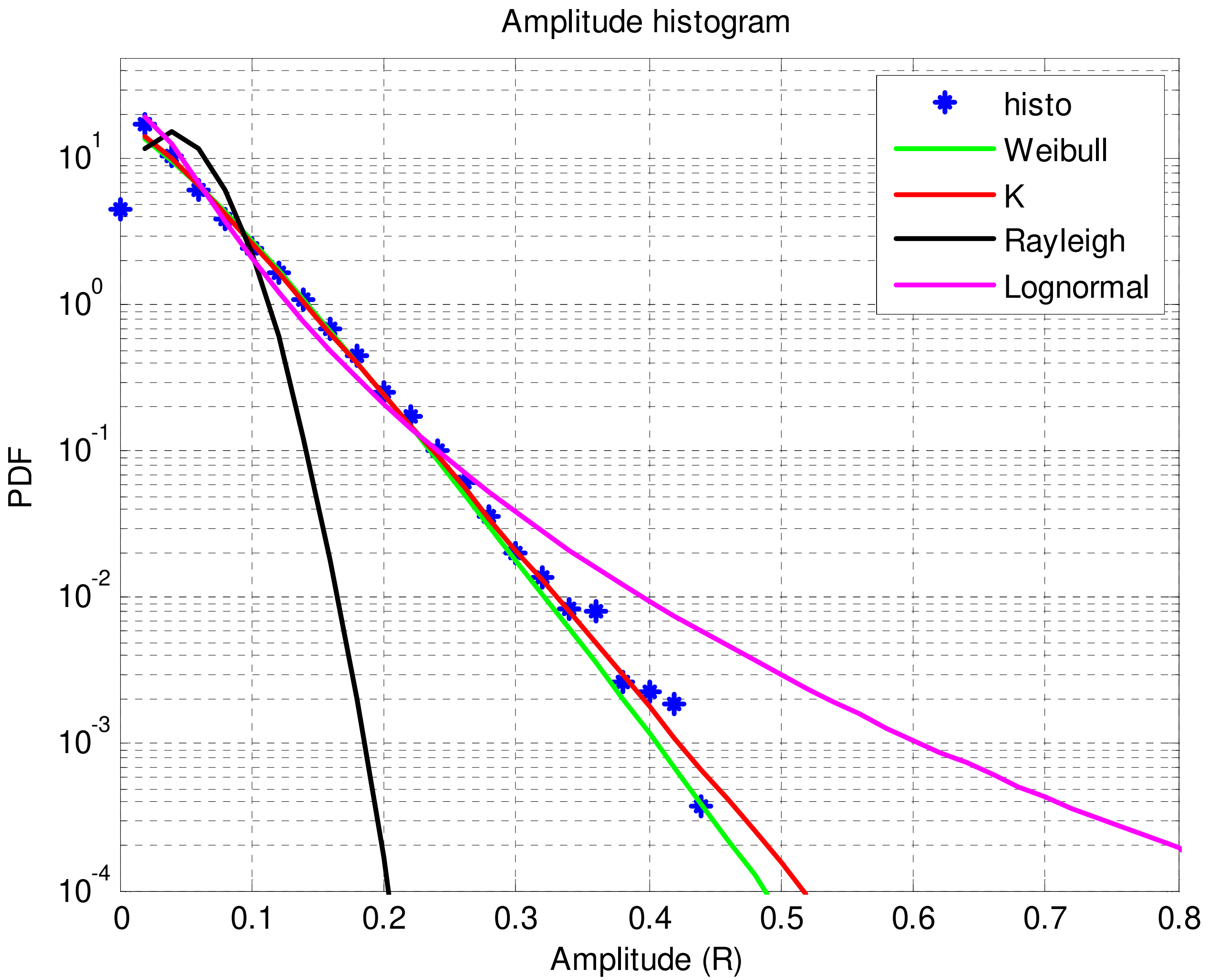
Applications: Scatter matrix estimation

- The **estimation of the correlation structure**, i.e. the scatter or covariance matrix, of a dataset is another common problem in many SP applications:
  - Adaptive detection in radar and sonar systems,
  - DOA estimation in array processing,
  - Principal Component Analysis (PCA),
  - Interference cancellation,
  - Portfolio optimization.
- Even if the data may come from disparate applications, they often share a non-Gaussian heavy-tailed statistical behavior (in radar and sonar applications, typically it is due to the high-resolution of the receiving sensor).

Real high-resolution sea clutter data



Heavy-tailed radar clutter



$$\text{Clutter amplitude:} \quad R = \left| X_I + jX_Q \right|$$

Scatter matrix estimation

- Estimating the scatter matrix of a set of non-Gaussian data is generally not a trivial task.

- The statistical characterization of non-Gaussian data requires additional parameters that generally have to be jointly estimated with the scatter matrix.

- The joint ML estimator of all unknown parameters often encounters computational difficulties and convergence (or even existence) issues.

- One could be led to use a simpler (mismatched) model.

A family of non-Gaussian distributions

- A popular family of non-Gaussian pdf's is the class of **Complex Elliptically Symmetric (CES)** distributions.

- Thanks to their flexibility, CES distributions represent a reliable data model in many areas such as radar, sonar, and communications [22].

- The complex Normal, Generalized Gaussian, *K*-distribution, complex *t*-distribution and all the compound-Gaussian pdf's belong to the CES class.

- The statistical behavior of high-resolution radar clutter can be accurately characterized by using the CES model ([24,25]).

Complex Elliptically Symmetric distributions

- A complex N-dimensional random vector $\mathbf{x}_m$ is Complex Elliptically Symmetric (CES) distributed if its pdf is [22]:

$$p_X(\mathbf{x}_m) = \left|\bar{\boldsymbol{\Sigma}}\right|^{-1} g\left(\left(\mathbf{x}_m - \boldsymbol{\gamma}\right)^H \bar{\boldsymbol{\Sigma}}^{-1}\left(\mathbf{x}_m - \boldsymbol{\gamma}\right)\right) \in CES_N\left(\boldsymbol{\gamma}, \bar{\boldsymbol{\Sigma}}, g\right)$$

- g is the density generator and γ the mean vector.

- $\bar{\boldsymbol{\Sigma}}$ is the full-rank scatter matrix, that is a scaled version of the covariance matrix $\bar{\mathbf{M}}$.

- A typical constraint is $\mathrm{tr}\left(\bar{\boldsymbol{\Sigma}}\right) = N$. As a consequence:

$$\bar{\mathbf{M}} \triangleq E\left\{(\mathbf{x}_m - \boldsymbol{\gamma})(\mathbf{x}_m - \boldsymbol{\gamma})^H\right\} = \sigma_X^2 \bar{\boldsymbol{\Sigma}}, \qquad \sigma_X^2 \triangleq \frac{E\left\{\mathbf{x}_m^H \mathbf{x}_m\right\}}{N} = \frac{\mathrm{tr}\left(\bar{\mathbf{M}}\right)}{N}$$

Estimation of the Scatter Matrix for CES data

- **Problem**: estimate the **scatter matrix** of a zero-mean CES distributed random vector in the presence of misspecified modeling, as described below.

- Set of *M* independent CES-distributed data: $\mathbf{x} = \{\mathbf{x}_m\}_{m=1}^{M} \in \mathbb{C}^N$

- True (unknown) data pdf: $p_X(\mathbf{x}_m) \in CES_N\left(\mathbf{0}, \bar{\boldsymbol{\Sigma}}, g\right)$

- Assumed data pdf: $f_X(\mathbf{x}_m; \boldsymbol{\theta}) \in CES_N\left(\mathbf{0}, \boldsymbol{\Sigma}, h_\zeta\right)$

$$\boldsymbol{\theta} = \left(\mathrm{vecs}(\boldsymbol{\Sigma})^T \quad \boldsymbol{\zeta}^T\right)^T$$

- The misspecification is in the choice of the density generator (which is parametrized by ζ).

A misspecified scenario

- A possible mismatched scenario in coherent radar [12]:
  - the **true** data pdf is a **complex *t*-distribution**;
  - the ML estimator of the scatter matrix (and of the statistical power) is derived under the **Gaussian model assumption**.

- The ML estimator of the scatter matrix under the Gaussian model assumption is the well-known **Sample Covariance Matrix (SCM)**.

1. Is the SCM a (misspecified) consistent estimator?
2. Is it efficient wrt the MCRB?
3. How large is its performance loss wrt the matched case?

Misspecified scatter matrix estimation

- **True model**: the *heavy-tailed* complex *t*-distribution.

$$p_X\left(\mathbf{x}_m \middle| \bar{\boldsymbol{\Sigma}}, \lambda, \eta\right) \triangleq \frac{1}{\pi^N \left|\bar{\boldsymbol{\Sigma}}\right|} \cdot \frac{\Gamma(N+\lambda)}{\Gamma(\lambda)} \left(\frac{\lambda}{\eta}\right)^{\lambda} \left(\frac{\lambda}{\eta} + \mathbf{x}_m^H \bar{\boldsymbol{\Sigma}}^{-1} \mathbf{x}_m\right)^{-(N+\lambda)}$$

- λ and η are the shape and scale parameters; they should be jointly estimated with the scatter matrix.

- The *statistical power* of $\mathbf{x}_m$ is: $\bar{\sigma}^2 = \dfrac{\lambda}{\eta(\lambda-1)}$

- The true, or "matched", parameter vector:

$$\boldsymbol{\tau} \triangleq [\mathrm{vecs}(\bar{\boldsymbol{\Sigma}})^T \quad \underbrace{\lambda \quad \eta}_{\zeta}]^T, \quad \mathrm{tr}(\bar{\boldsymbol{\Sigma}}) = N$$

Misspecified scatter matrix estimation

- **Assumed model**: the complex Gaussian distribution.

$$f_X\left(\mathbf{x}_m \middle| \boldsymbol{\theta}\right) \triangleq f_X\left(\mathbf{x}_m \middle| \boldsymbol{\Sigma}, \sigma^2\right) = \frac{1}{\left(\pi\sigma^2\right)^N \left|\boldsymbol{\Sigma}\right|} \exp\left(-\frac{\mathbf{x}_m^H \boldsymbol{\Sigma}^{-1} \mathbf{x}_m}{\sigma^2}\right)$$

- The "assumed" parameter vector to be estimated is:

$$\boldsymbol{\theta} = [\mathrm{vecs}(\boldsymbol{\Sigma})^T \quad \sigma^2]^T$$

- under the <u>constraint</u>:

$$\mathbf{f}\left(\boldsymbol{\theta}\right) = \mathbf{0} \quad \Rightarrow \quad \mathrm{tr}\left(\boldsymbol{\Sigma}\right) - N = 0$$

- Without such (or similar) constraint, $\boldsymbol{\Sigma}$ and σ^2 cannot be jointly estimated $\rightarrow$ there is an ambiguity.

Misspecified scatter matrix estimation

- The **constrained MML (CMML) estimator** is derived as:

$$\hat{\boldsymbol{\theta}}_{CMML}(\mathbf{x}) \triangleq \underset{\boldsymbol{\theta}\in\Theta}{\arg\max} \ln f_X(\mathbf{x};\boldsymbol{\theta})$$

→ assumed complex Gaussian distribution

$$= \underset{\boldsymbol{\theta}\in\Theta}{\arg\max} \sum_{m=1}^{M} \ln f_X(\mathbf{x}_m;\boldsymbol{\theta}), \qquad \mathbf{x}_m \sim p_X(\mathbf{x}_m)$$

true (unknown) complex t-distribution

- The closed form expression is:

$$\begin{cases} \hat{\boldsymbol{\Sigma}}_{CMML} = \dfrac{N}{Tr\{\mathbf{SCM}\}}\mathbf{SCM} = \dfrac{N}{Tr\left\{\dfrac{1}{M}\sum_{m=1}^{M}\mathbf{x}_m\mathbf{x}_m^H\right\}} \cdot \dfrac{1}{M}\sum_{m=1}^{M}\mathbf{x}_m\mathbf{x}_m^H \\ \hat{\sigma}^2_{CMML} = \dfrac{1}{M}\sum_{m=1}^{M}\dfrac{\mathbf{x}_m^H\hat{\boldsymbol{\Sigma}}_{CMML}^{-1}\mathbf{x}_m}{N} \end{cases}$$

Misspecified scatter matrix estimation

- The **CMML estimator** under the Gaussian assumption:

$$\begin{cases} \hat{\boldsymbol{\Sigma}}_{CMML} = \dfrac{N}{\sum_{m=1}^{M} \mathbf{x}_m^H \mathbf{x}_m} \sum_{m=1}^{M} \mathbf{x}_m \mathbf{x}_m^H \\ \hat{\sigma}_{CMML}^2 = \dfrac{1}{NM} \sum_{m=1}^{M} \mathbf{x}_m^H \hat{\boldsymbol{\Sigma}}_{CMML}^{-1} \mathbf{x}_m \end{cases}$$

Which is the convergence point of the CMML estimator?

$$\hat{\boldsymbol{\theta}}_{CMML}(\mathbf{x}) \underset{M\to\infty}{\overset{a.s.}{\to}} \boldsymbol{\theta}_0 = ?$$

$$\hat{\boldsymbol{\theta}}_{CMML} = \left[\mathrm{vecs}\left(\hat{\boldsymbol{\Sigma}}_{CMML}\right)^T \quad \hat{\sigma}_{CMML}^2 \right]^T$$

Misspecified scatter matrix estimation

- The pseudo-true parameter vector is:

$$\boldsymbol{\theta}_0 = \arg\min_{\boldsymbol{\theta}\in\tilde{\Theta}} \left\{ -E_p \left\{ \ln f_X (\mathbf{x}_m; \boldsymbol{\theta}) \right\} \right\}$$

- For the case study at hand, it can be shown that:

$$\boldsymbol{\theta}_0 = [\text{vecs}(\bar{\boldsymbol{\Sigma}})^T \quad \bar{\sigma}^2]^T \qquad \text{and} \qquad \det\left(\mathbf{A}_{\boldsymbol{\theta}_0}\right) \neq 0$$

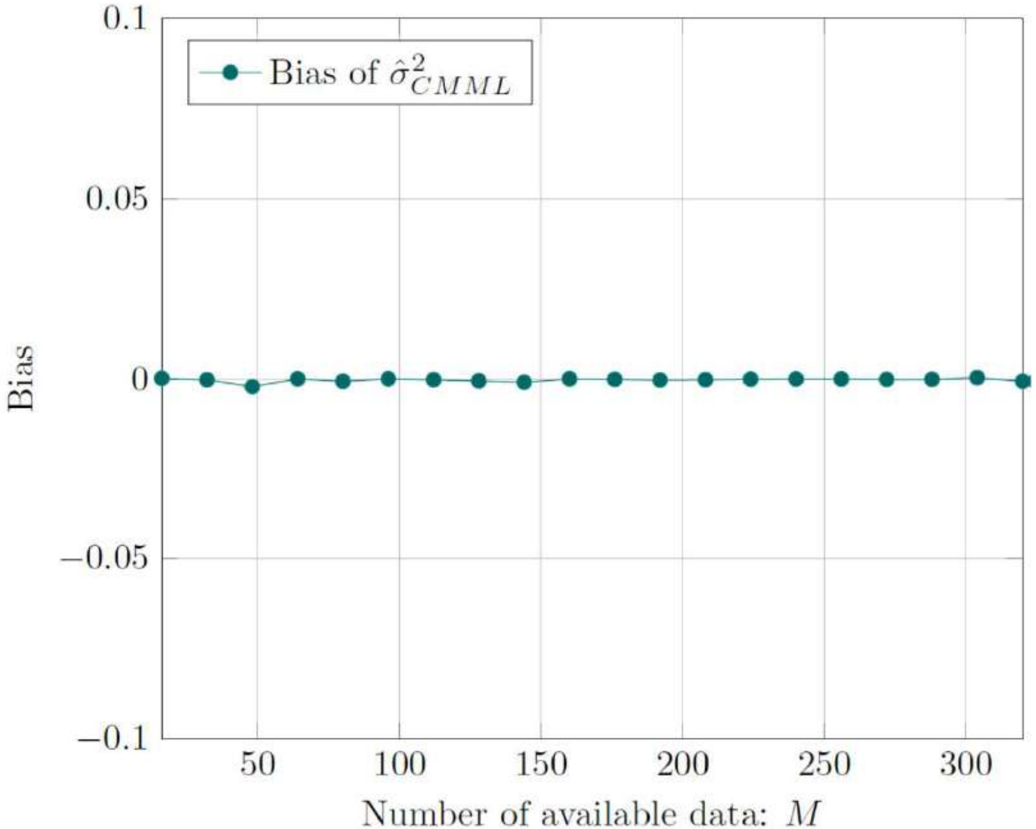
The **CMMLE** converges ***a.s.*** to:

$$\hat{\sigma}^2_{CMML}(\mathbf{x}) \underset{M\to\infty}{\overset{a.s.}{\rightarrow}} \bar{\sigma}^2 = \frac{\lambda}{\eta(\lambda-1)}, \quad \hat{\boldsymbol{\Sigma}}_{CMML}(\mathbf{x}) \underset{M\to\infty}{\overset{a.s.}{\rightarrow}} \bar{\boldsymbol{\Sigma}}$$

Hence, it provides **consistent** estimates for both the statistical power and the scatter matrix.

Misspecified scatter matrix estimation

- Bias of the estimator $\hat{\sigma}^2_{CMML}$: $b_{\hat{\sigma}^2_{CMML}} \triangleq E\left\{\hat{\sigma}^2_{CMML} - \bar{\sigma}^2\right\}$



Simulation parameters:

$$[\bar{\mathbf{\Sigma}}]_{i,j} = \rho^{|i-j|}$$

$$\rho = 0.8$$

$$N = 16$$

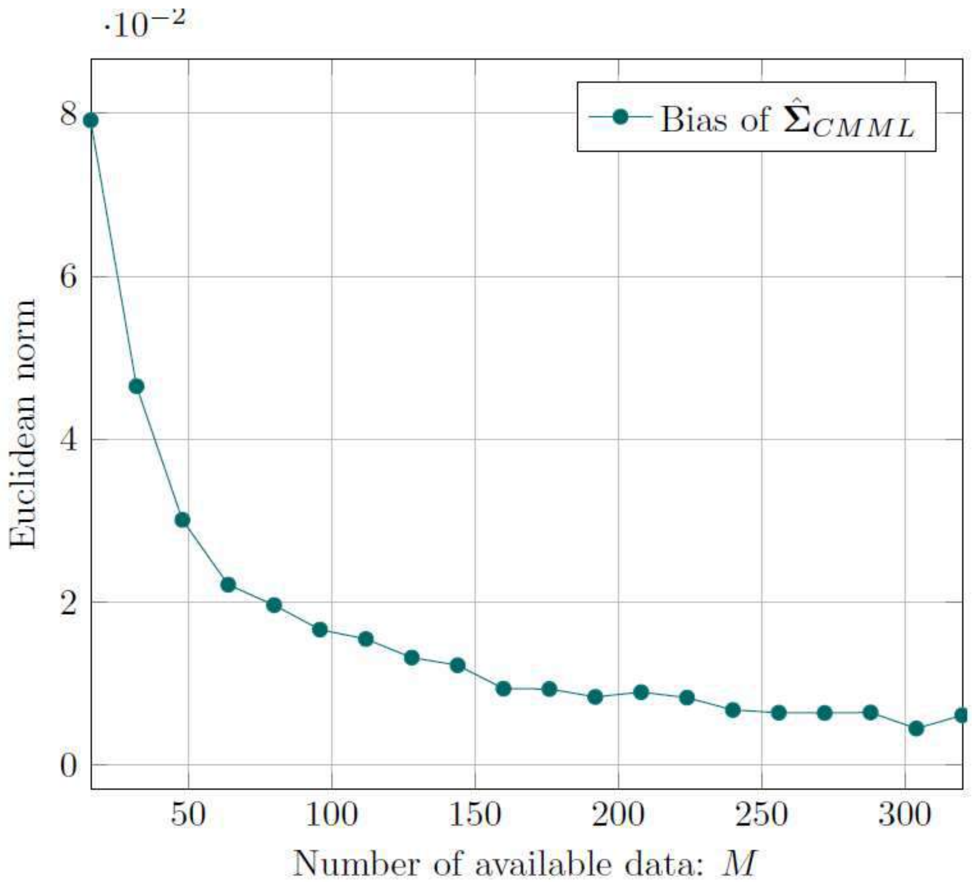
$$\lambda = 3$$

$$\eta = 1$$

$$\bar{\sigma}^2 = \frac{\lambda}{\eta(\lambda - 1)} = \frac{3}{2}$$

Misspecified scatter matrix estimation

- Bias of the estimator $\hat{\boldsymbol{\Sigma}}_{CMML}$: $b_{\hat{\boldsymbol{\Sigma}}_{CMML}} \triangleq \left\| E\left\{ \text{vecs}(\hat{\boldsymbol{\Sigma}}_{CMML} - \bar{\boldsymbol{\Sigma}}) \right\} \right\|_2$



Simulation parameters:

$$[\bar{\boldsymbol{\Sigma}}]_{i,j} = \rho^{|i-j|}$$

$$\rho = 0.8$$

$$N = 16$$

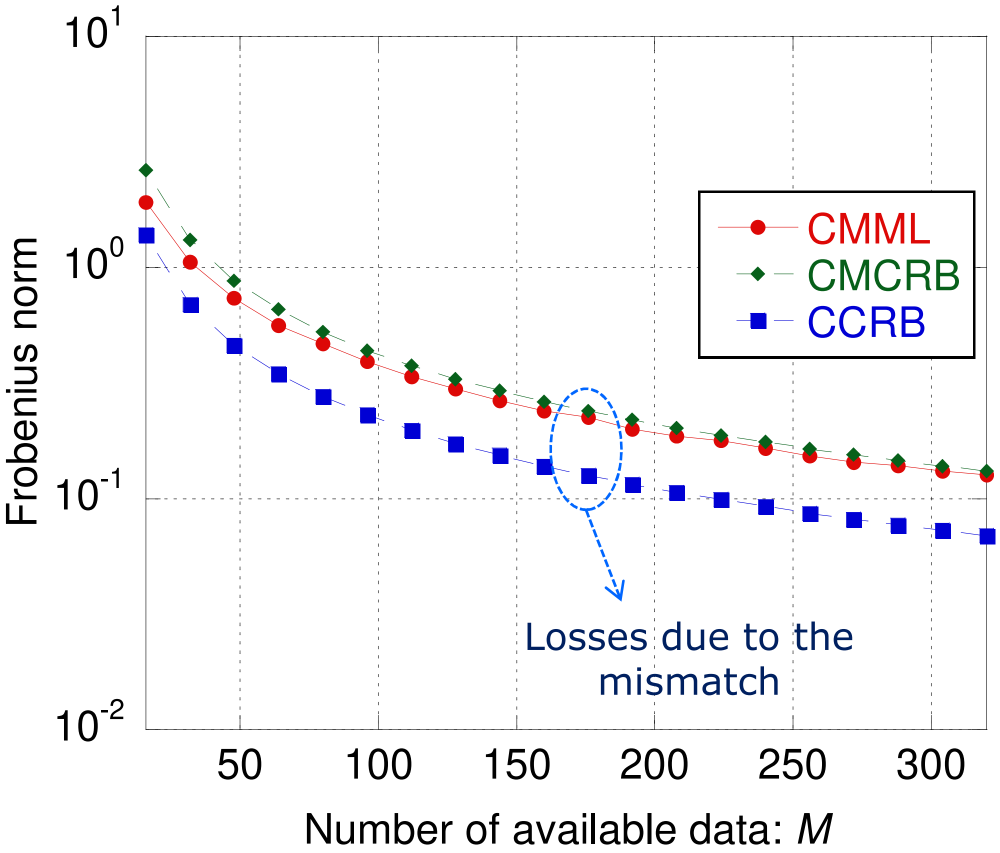
$$\lambda = 3$$

$$\eta = 1$$

$$\bar{\sigma}^2 = \frac{\lambda}{\eta(\lambda - 1)} = \frac{3}{2}$$

Misspecified scatter matrix estimation

$$\varepsilon_{\hat{\Sigma}_{CMML}} \triangleq \left\| E\left\{ \mathrm{vecs}(\hat{\boldsymbol{\Sigma}}_{CMML} - \bar{\boldsymbol{\Sigma}})\mathrm{vecs}(\hat{\boldsymbol{\Sigma}}_{CMML} - \bar{\boldsymbol{\Sigma}})^T \right\} \right\|_F$$



Simulation parameters:

$$[\bar{\boldsymbol{\Sigma}}]_{i,j} = \rho^{|i-j|}$$

$$\rho = 0.8$$

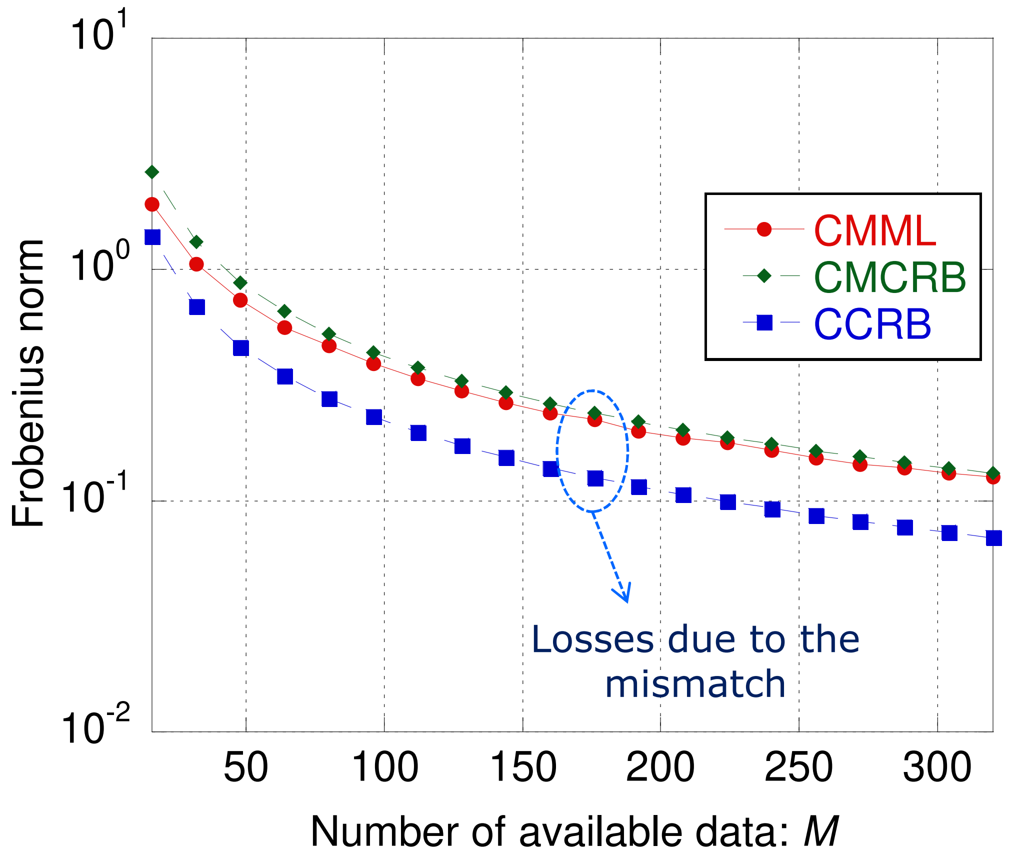
$$N = 16$$

$$\lambda = 3$$

$$\eta = 1$$

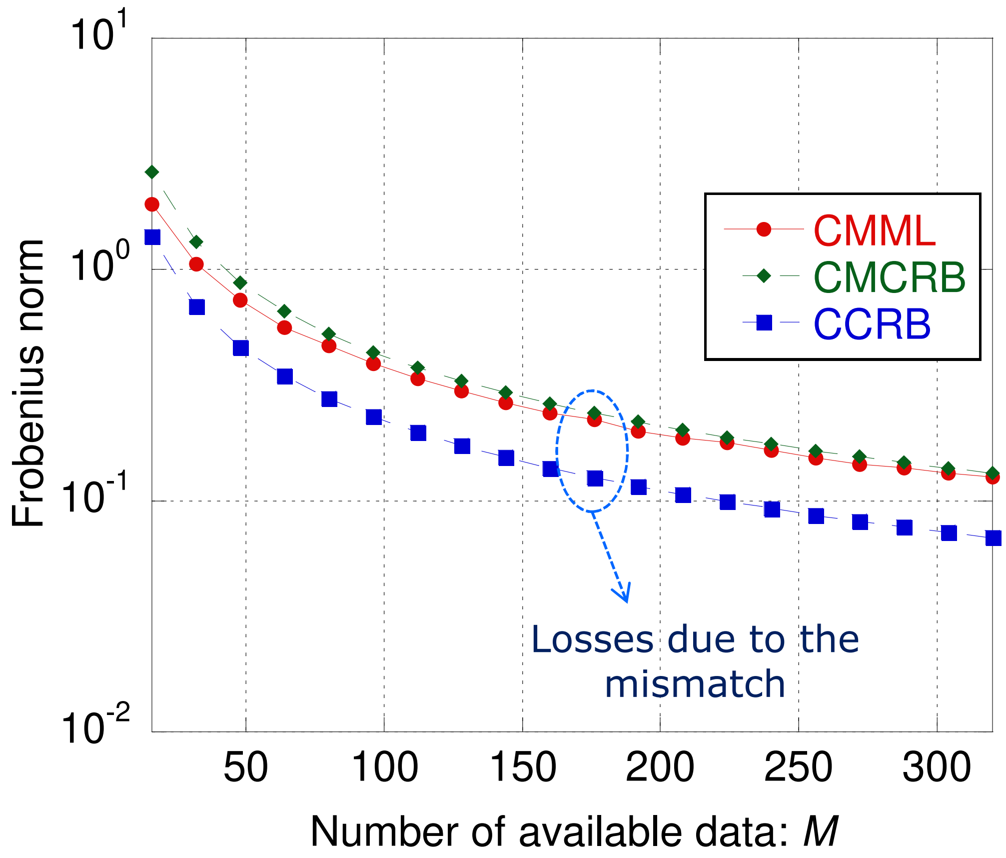
$$\bar{\sigma}^2 = \frac{\lambda}{\eta(\lambda - 1)} = \frac{3}{2}$$

Misspecified scatter matrix estimation



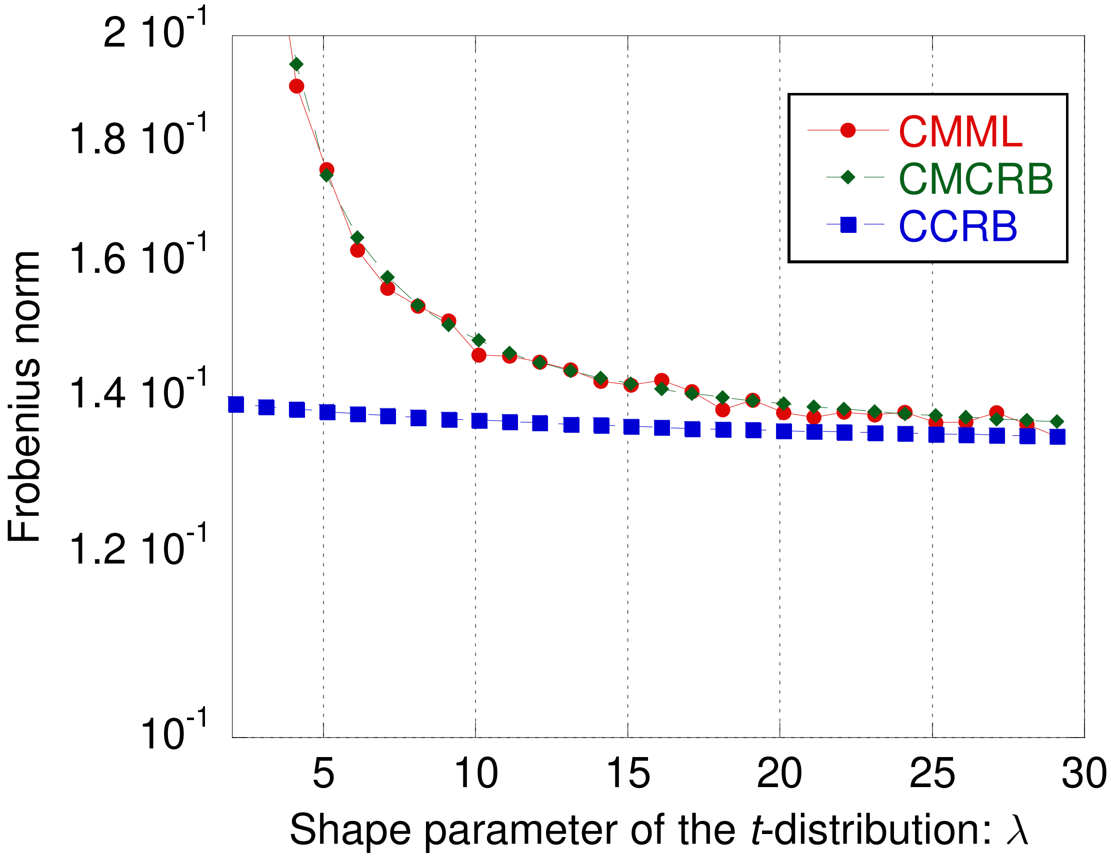
- **CCRB** is the **constrained** "matched" **CRB** [30,31,32] on the joint estimation of $\bar{\Sigma}$, λ, and η (for t-distributed data).

Misspecified scatter matrix estimation



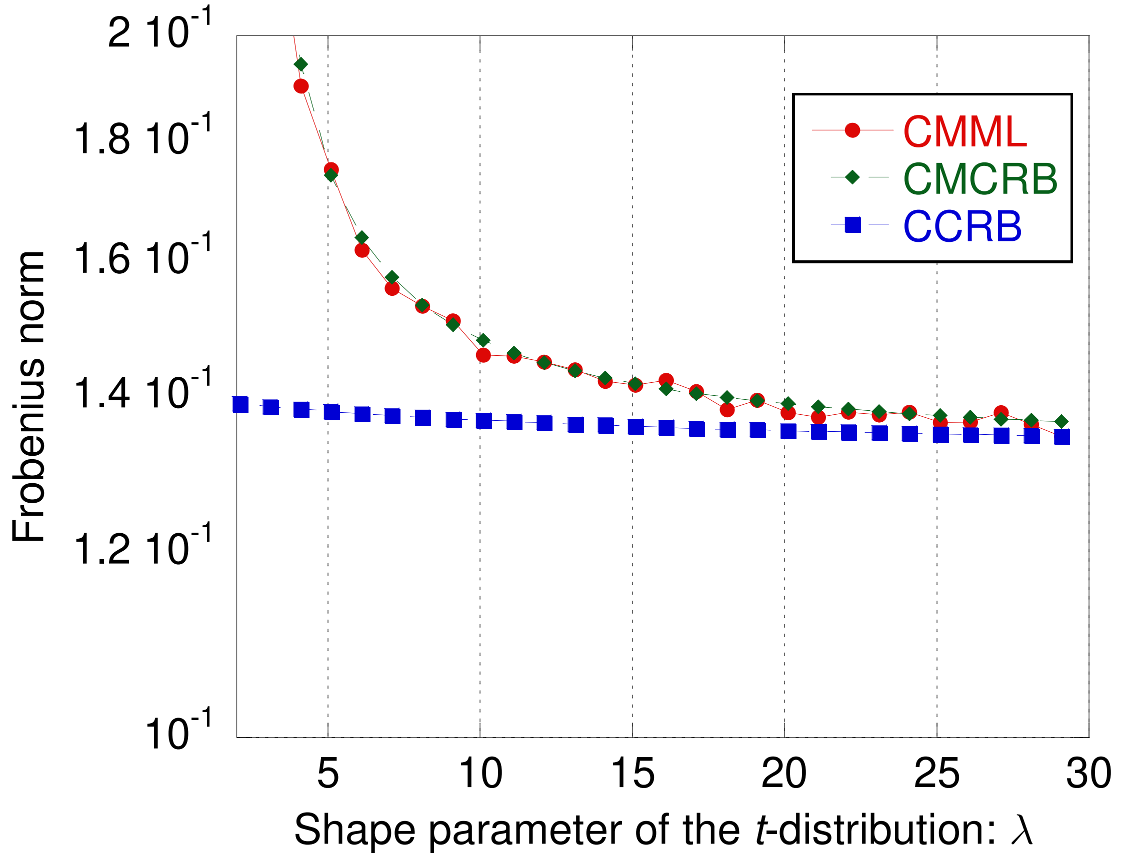
- **CMCRB** is the **constrained MCRB** [20] on the joint estimation of $\bar{\Sigma}$ and $\bar{\sigma}^2$(under model misspecification).

Misspecified scatter matrix estimation



- When $\lambda \rightarrow 0$ (extremely spiky data), the estimation losses due to model mismatching rapidly increase.

Misspecified scatter matrix estimation



- When $\lambda \to \infty$ the data tend to be Gaussian distributed and the MSE of the **CMMLE**, the **CMCRB**, and the **CCRB** tend to coincide.

Concluding remarks for Part I

- We summarized the fundamental concepts about lower bounds and efficient estimators in the presence of model misspecification.

- The MML estimator is asymptotically MS-unbiased and its error covariance matrix asymptotically equates the MCRB.

- We showed how to apply these theoretical findings to two well-known problems:
    1. Direction of Arrivals (DOAs) estimation with an array of antennas;
    2. Estimation of the the disturbance scatter matrix in complex *t*-distributed data.

Main references

[1] P. J. Huber, "The behavior of Maximum Likelihood Estimates under Nonstandard Conditions," *Proc. of the Fifth Berkeley Symposium in Mathematical Statistics and Probability*. Berkley: University of California Press, 1967.

[2] H. Akaike, "Information theory and an extension of the likelihood principle", *Proc. of 2nd International Symposium of Information Theory*, pp. 267-281, 1972.

[3] H. White, "Maximum likelihood estimation of misspecified models", *Econometrica* vol. 50, pp. 1-25, January 1982.

[4] Q. H. Vuong, "Cramér-Rao bounds for misspecified models", *Working paper 652, Division of the Humanities and Social Sciences ,* Caltech, October 1986.

[5] R. H. Berk, "Limiting behaviour of posterior distributions when the model is incorrect", *Ann. Math. Statist.,* vol. 37, pp. 51-58, 1966.

[6] O. Bunke, X. Milhaud, "Asymptotic Behavior of Bayes Estimates Under Possibly Incorrect Models", *The Annals of Statistics*, vol. 26, No. 2, pp. 617–644, 1998.

[7] C. D. Richmond, P. Basu, "Bayesian framework and radar: on misspecified bounds and radar-communication cooperation," *IEEE Statistical Signal Processing Workshop 2016 (SSP)*, Palma de Mallorca, Spain, 26–29 June 2016.

[8] C. D. Richmond, L. L. Horowitz, "Parameter Bounds on Estimation Accuracy Under Model Misspecification," *IEEE Trans. on Signal Processing*, vol. 63, No. 9, pp. 2263-2278, May 1, 2015.

Main references

[9] C. Ren, M. N. El Korso, J. Galy, E. Chaumette, P. Larzabal, A. Renaux, “Performances bounds under misspecification model for MIMO radar application” *European Signal Process. Conf. (EUSIPCO)*, Nice, France, pp. 514-518, 2015.

[10] M. S. Greco, S. Fortunati, F. Gini, “Maximum likelihood covariance matrix estimation for complex elliptically symmetric distributions under mismatched conditions,” *Signal Processing*, vol. 104, pp. 381-386, November 2014.

[11] S. Fortunati, F. Gini, M. S. Greco, “The Misspecified Cramér-Rao Bound and its Application to the Scatter Matrix estimation in Complex Elliptically Symmetric distributions,” *IEEE Trans. Signal Process*ing, Vol. 64, No. 9, pp. 2387-2399, 2016.

[12] S. Fortunati, F. Gini, M. S. Greco, “Matched, mismatched and robust scatter matrix estimation and hypothesis testing in complex t-distributed data”, *EURASIP Journal on Advances in Signal Processing* (2016) 2016:123.

[13] P. A. Parker and C. D. Richmond, “Methods and Bounds for Waveform Parameter Estimation with a Misspecified Model,” *Conf. on Signals, Systems, and Computers (Asilomar)*, Pacific Grove, CA, pp. 1702-1706, November 2015.

[14] A. Gusi-Amigó, P. Closas, A. Mallat and L. Vandendorpe, “Ziv-Zakai lower bound for UWB based TOA estimation with unknown interference,” *IEEE ICASSP 2014, Florence*, 2014, pp. 6504-6508.

[15] H. Krim, M. Viberg, “Two decades of array signal processing research: the parametric approach,” *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67-94, July 1996.

Main references

[16] B. Friedlander, "Sensitivity analysis of the maximum likelihood direction-finding algorithm," *IEEE Trans. on Aerospace and Electronics Systems*, Vol. 26, No. 6, pp. 953-968, Nov 1990.

[17] H. L. Van Trees, K. L. Bell, and Z. Tian, *Detection, Estimation and Modulation Theory:* Vol.1, 2nd ed. Hoboken: Wiley, 2013.

[18] Y. Rockah and P. Schultheiss, "Array shape calibration using sources in unknown locations-Part I: Far-field sources," *IEEE Trans on Acoust., Speech, Signal Process*., Vol. 35, No. 3, pp. 286-299, Mar 1987.

[19] M. Pardini, F. Lombardini and F. Gini, "The Hybrid Cramér–Rao Bound on Broadside DOA Estimation of Extended Sources in Presence of Array Errors," *IEEE Trans. on Signal Process.*, Vol. 56, No. 4, pp. 1726-1730, April 2008.

[20] S. Fortunati, F. Gini, M. S. Greco, "The Constrained Misspecified Cramér-Rao Bound," *IEEE Signal Processing Letters*, Vol. 23, No. 5, pp. 718-721, May 2016.

[21] S. Fortunati, "Misspecified Cramér-Rao Bounds for Complex Unconstrained and Constrained Parameters," in *25th European Signal Processing Conference (EUSIPCO)*, Kos, Greece, 28 Aug. – 2 Sept. 2017.

[22] E. Ollila, D. E. Tyler, V. Koivunen, and H. V. Poor, "Complex elliptically symmetric distributions: Survey, new results and applications," *IEEE Transactions on Signal Processing*, Vol. 60, No. 11, pp. 5597–5625, 2012.

Main references

[23] A. Mennad, S. Fortunati, M. N. El Korso, A. Younsi, A. M. Zoubir, A. Renaux, "Slepian-Bangs-type formulas and the related Misspecified Cramér-Rao Bounds for Complex Elliptically Symmetric Distributions", *Signal Processing,* Vol. 142, pp. 320-329, Jan. 2018.

[24] A. Younsi, M. Greco, F. Gini, and A. Zoubir, "Performance of the adaptive generalised matched subspace constant false alarm rate detector in non-Gaussian noise: An experimental analysis," *IET Radar Sonar Navig*., vol. 3, no. 3, pp. 195–202, 2009..

[25] K. J. Sangston, F. Gini, M. Greco, "Coherent radar detection in heavy-tailed compound-Gaussian clutter", *IEEE Trans. on Aerospace and Electronic Syst*ems, Vol. 48, No. 1, pp. 64-77, January 2012.

[26] C. D. Richmond, "On Constraints in Parameter Estimation and Model Misspecification," *International Conference on Information Fusion (FUSION)*, Cambridge, 2018, pp. 1080-1085.

[27] M. Pajovic, "Misspecified Bayesian Cramér-Rao Bound for Sparse Bayesian," *IEEE Statistical Signal Processing Workshop (SSP)*, Freiburg, 2018, pp. 263-267.

[28] X. Zhou and S. Kay, "A Robust Spectral Estimator with Application to a Noise Corrupted Process," *IEEE Transactions on Signal Processing*, vol. 67, no. 8, pp. 2107-2114, 15 April15, 2019.

Main references

[29] P. Wang, T. Koike-Akino, M. Pajovic, P. V. Orlik, W. Tsujita, and F. Gini, "Misspecified CRB on Parameter Estimation for a Coupled Mixture of Polynomial Phase and Sinusoidal FM Signals," *ICASSP 2019,* Brighton, UK, 2019, pp. 5302-5306.

[30] J. D. Gorman and A. O. Hero, "Lower bounds for parametric estimation with constraints," *IEEE Trans. Inf. Theory*, vol. 6, no. 6, pp. 1285–1301, Nov. 1990.

[31] T. J. Moore, R. J. Kozick, and B. M. Sadler, "The constrained Cramér-Rao bound from the perspective of fitting a model," *IEEE Signal Process. Lett.*, vol. 14, no. 8, pp. 564–567, Aug. 2007.

[32] P. Stoica and B. C. Ng, "On the Cramer–Rao bound under parametric constraints," *IEEE Signal Process. Lett.*, vol. 5, no. 7, pp. 177–179, July 1998.