AN IDENTIFIABILITY CRITERION IN THE PRESENCE OF RANDOM NUISANCE PARAMETERS

Stefano Fortunati[*,1], *Fulvio Gini*[*], *Maria Greco*[*], *Alfonso Farina*[#], *Antonio Graziano*[#], *Sofia Giompapa*[#]

[*]Dipartimento di Ingegneria dell'Informazione, University of Pisa,
via G. Caruso 16, 56122, Italy.
{stefano.fortunati, f.gini, m.greco}@iet.unipi.it
[#]SELEX Sistemi Integrati
Via Tiburtina, km 12.4, 00123 Roma, Italy
{afarina, agraziano, sgiompapa}@selex-si.com

ABSTRACT

This paper concerns with the identifiability of an unknown deterministic vector in the presence of random nuisance parameters. In these cases, the classical definition of identifiability, which requires calculation of the Fisher Information Matrix (FIM) and of its rank, is often difficult or impossible to be implemented. Instead, the Modified FIM (MFIM) can be usually computed. We generalize the main results on parameter identifiability to take the presence of random nuisance parameters into account. We provide an alternative definition of identifiability that can be always applied also in the presence of nuisance parameters and we investigate the relationships between the classical and the new identifiability conditions. Finally, the new definition of identifiability is applied to a common estimation problem in netted radar systems: the relative grid-locking problem.

Index Terms— Identifiability, nuisance parameters, Kullback-Leibler divergence, Modified FIM.

1. INTRODUCTION

The identifiability problem concerns with the ability of drawing inferences from observed data to an underlying theoretical structure [1], [2], [3]. This is equivalent to saying that different structures may generate different probability distributions of the observable data in order to make the structures "observable". Our attention is focused on the parametric models [4], i.e. such models in which every structure can be represented by a vector in $\mathbb{R}^m$. The classical definition of identifiability requires the calculation of the rank of the Fisher Information Matrix (FIM) [2]: if the FIM is a full rank matrix, then the structure (or, equivalently its corresponding parameter vector) is identifiable. However, in many practical applications, the data model is affected by additional random parameters whose estimation is not strictly required, the so-called nuisance parameters [5]. In these cases, the evaluation of the FIM is often difficult or impossible to be evaluated in closed form. To overcome this analytical problem, the modified FIM (MFIM) has been introduced [6], [7]. The aim of this paper is to generalize the classical identifiability condition to take the presence of random nuisance parameters into account. In particular, a new definition of identifiability, based on the rank of the MFIM, is provided. This alternative definition is weaker than the classical one in the sense that some structures might be classified as identifiable when they are not according to the classical condition. On the other hand, if a structure is not identifiable under the proposed new condition, then it is not identifiable under the classical condition as well. We show the relationships between the new identifiability condition and the MFIM. Finally, we apply our results to the relative grid-locking problem which shows up in netted radar systems (see [8],[9] and references therein).

The rest of the paper is organized as follows: in Sect. 2, we provide a description of the identifiability problem in its classical formulation. In Sect. 3, we generalize previous results to take into account the presence of random nuisance parameters. A new definition of identifiability is provided and its relationship with the classical one investigated. Finally, in Sect. 4, we describe an application to the relative grid-locking problem. Conclusions are collected in Sect. 5.

2. GENERAL FORMULATION OF THE IDENTIFICABILITY PROBLEM

2.1. Some preliminary definition

Let $\mathbf{x} \in \mathbb{R}^n$ be a n-dimensional random vector, representing the outcome of some random experiment, whose probability density function (pdf) is known to belong to a family $\mathcal{F}$. A *structure* T is a set of hypotheses which implies a unique pdf in $\mathcal{F}$ for $\mathbf{x}$. Such pdf is indicated with

[1] The work of S. Fortunati has been supported by SESM-Finmeccanica and SELEX Sistemi Integrati.

$p(\mathbf{x};T) \in \mathcal{F}$ [1], [3]. The set of all the a priori possible structures is called a *model* and is denoted by $\mathcal{M}$. By definition, there exist a unique pdf associated with each structure in $\mathcal{M}$.

Definition 1: *Two structures T_0 and T_1 in $\mathcal{M}$ are said to be observationally equivalent if they imply the same pdf for the observable random vector* $\mathbf{x}$. *The structure T_0 is otherwise said to be identifiable if there is no other structure in $\mathcal{M}$ which is observationally equivalent.*

We assume that the pdf of the random vector $\mathbf{x}$ has a parametric representation, i.e. we assume that every structure T is described by an m-dimensional vector $\mathbf{\Phi}$ and that the model is described by a open subset $\Omega \subseteq \mathbb{R}^m$. It is possible to associate with each $\mathbf{\Phi}$ in Ω a continuous pdf $p(\mathbf{x};\mathbf{\Phi}) \in \mathcal{F}$ which is perfectly known except for the values of the parameter vector $\mathbf{\Phi}$.

Definition 2: *Two parameter vectors $\mathbf{\Phi}_0$ and $\mathbf{\Phi}_1$ (relative to two structures T_0 and T_1) are said to be observationally equivalent if $p(\mathbf{x};\mathbf{\Phi}_0)=p(\mathbf{x};\mathbf{\Phi}_1)$ for all $\mathbf{x} \in \mathbb{R}^n$. $\mathbf{\Phi}_0$ is otherwise said to be identifiable if there is no other $\mathbf{\Phi}$ in Ω which is observationally equivalent.*

Since the set of the structures is an open subset of $\mathbb{R}^m$ then it is possible to endow it with the same topological structure of $\mathbb{R}^m$. This allows us to consider the concept of local identifiability:

Definition 3: *A parameter vector $\mathbf{\Phi}_0$ is said to be locally identifiable if there exists an open neighborhood of $\mathbf{\Phi}_0$ containing no other $\mathbf{\Phi}$ in Ω which is observationally equivalent.*

To highlight the difference between the Definitions 2 and 3, in the following we indicate as *global* the identifiability in Definition 2 and as *local* the identifiability in Definition 3. In the following, we assume satisfied some regularity conditions on $p(\mathbf{x};\mathbf{\Phi}_0)$ [1].

2.2. A general identifiability criterion

In [2], a general criterion, based on the Kullback-Liebler (KL) divergence, for the identifiability of parameter vectors is proposed. Here we report only the main facts. All the proofs can be found in [2]. First of all, we recall the definition of the KL divergence [10]:

Definition 4: *Let $p(\mathbf{x};\mathbf{\Phi})$ and $p(\mathbf{x};\mathbf{\Phi}_0)$ be two parametric pdfs for all $\mathbf{\Phi} \in \Omega$. The scalar function of the vector variable $\mathbf{\Phi}$, $H(\mathbf{\Phi};\mathbf{\Phi}_0)$, defined as*:

$$H(\mathbf{\Phi};\mathbf{\Phi}_0) \triangleq E_{\mathbf{x}}\left\{\ln\frac{p(\mathbf{x};\mathbf{\Phi})}{p(\mathbf{x};\mathbf{\Phi}_0)}\right\} = \int \ln\frac{p(\mathbf{x};\mathbf{\Phi})}{p(\mathbf{x};\mathbf{\Phi}_0)}\, p(\mathbf{x};\mathbf{\Phi}_0)\, d\mathbf{x} \tag{1}$$

is called KL divergence between $p(\mathbf{x};\mathbf{\Phi})$ and $p(\mathbf{x};\mathbf{\Phi}_0)$.

One of the most important theorems on the KL divergence is

Theorem 1: *Let $p(\mathbf{x};\mathbf{\Phi})$ and $p(\mathbf{x};\mathbf{\Phi}_0)$ be two parametric pdfs. If $p(\mathbf{x};\mathbf{\Phi})=p(\mathbf{x};\mathbf{\Phi}_0)$ for all $\mathbf{x} \in \mathbb{R}^n$, then $H(\mathbf{\Phi};\mathbf{\Phi}_0)=0$. Otherwise, if $H(\mathbf{\Phi};\mathbf{\Phi}_0)$ is finite, $H(\mathbf{\Phi};\mathbf{\Phi}_0)<0$.*

In view of Definitions 2, 3, and 4, the link between the KL divergence and the identifiability of a parameter vector is given by the following corollary:

Corollary 1: *Let $p(\mathbf{x};\mathbf{\Phi})$ and $p(\mathbf{x};\mathbf{\Phi}_0)$ be two parametric pdfs for all $\mathbf{\Phi} \in \Omega$. Then the parameter vector $\mathbf{\Phi}_0$ is globally identifiable if and only if the equation $H(\mathbf{\Phi};\mathbf{\Phi}_0)=0$ has, as solution in Ω, only $\mathbf{\Phi}=\mathbf{\Phi}_0$. It is locally identifiable if and only if $\mathbf{\Phi}=\mathbf{\Phi}_0$ is the only solution in some open neighborhood of $\mathbf{\Phi}_0$.*

It can be noted also that the identifiability condition is closely related to the maximum of the $H(\mathbf{\Phi};\mathbf{\Phi}_0)$. In fact, from Theorem 1 it follows that: if the maximum of $H(\mathbf{\Phi};\mathbf{\Phi}_0)$ is global and attained at $\mathbf{\Phi}=\mathbf{\Phi}_0$, then $\mathbf{\Phi}_0$ is globally identifiable, whereas, if there exists an open neighborhood of $\mathbf{\Phi}_0$ with a local maximum in $\mathbf{\Phi}_0$, then $\mathbf{\Phi}_0$ is locally identifiable. Such consideration suggests another general identification criterion that we provide in the following corollary (the proof can be found in [2]).

Corollary 2: *Let $p(\mathbf{x};\mathbf{\Phi})$ and $p(\mathbf{x};\mathbf{\Phi}_0)$ be two parametric pdf's for all $\mathbf{\Phi} \in \Omega$. Then, the parameter vector $\mathbf{\Phi}_0$ is locally identifiable if and only if the Hessian matrix* $\mathbf{H}$ *of $H(\mathbf{\Phi};\mathbf{\Phi}_0)$ evaluated at $\mathbf{\Phi}_0$, i.e. $\boldsymbol{H}(H)(\mathbf{\Phi}_0)$, is a negative definite matrix. Moreover, it can be shown that:*

$$\begin{aligned}\left[\mathbf{H}(H)(\mathbf{\Phi}_0)\right]_{ij} &= -E\left\{\frac{\partial}{\partial \Phi_i}\ln p(\mathbf{x};\mathbf{\Phi})\Big|_{\mathbf{\Phi}=\mathbf{\Phi}_0}\frac{\partial}{\partial \Phi_j}\ln p(\mathbf{x};\mathbf{\Phi})\Big|_{\mathbf{\Phi}=\mathbf{\Phi}_0}\right\} \\ &= E\left\{\frac{\partial^2}{\partial \Phi_i \partial \Phi_j}\ln p(\mathbf{x};\mathbf{\Phi})\Big|_{\mathbf{\Phi}=\mathbf{\Phi}_0}\right\} = -\left[\mathbf{I}(\mathbf{\Phi}_0)\right]_{ij},\end{aligned} \tag{2}$$

where $\mathbf{I}(\mathbf{\Phi}_0)$ is the Fisher Information Matrix (FIM).

Taking into account eq. (2), the following Corollary can be finally derived [2]:

Corollary 3: *Let $p(\mathbf{x};\mathbf{\Phi}_0)$ be a parametric pdf. Then the parameter vector $\mathbf{\Phi}_0$ is locally identifiable if and only if the Fisher Information Matrix $\mathbf{I}(\mathbf{\Phi}_0)$ is a full rank matrix.*

3. IDENTIFIABILITY IN PRESENCE OF RANDOM NUISANCE PARAMETERS

In practical applications, a wide class of estimation problem involves the so-called nuisance parameters, i.e. random parameters that affect the data model, whose estimation is not strictly required and that are known only through their statistical distribution. As before, let $\mathbf{x} \in \mathbb{R}^n$ be a n-dimensional random vector, representing the outcome of some random experiment, let $\mathbf{a} \in \mathbb{R}^l$ be the l-dimensional random vector of nuisance parameters and let $p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_0)$ be the joint pdf of the random vectors $\mathbf{x}$ and $\mathbf{a}$ parameterized by the deterministic vector $\mathbf{\Phi}_0$ to be estimated. Such pdf is assumed perfectly known. In the rest of the paper, we assume verified the following:

Assumption 1: *The pdf of the nuisance parameters* $p(\mathbf{a})$ *does not depend on the parameter vector* $\mathbf{\Phi}_0$. *Then, the joint pdf* $p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_0)$ *can be always factorized as* $p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_0)=$ $=p(\mathbf{x}|\mathbf{a};\mathbf{\Phi}_0)p(\mathbf{a};\mathbf{\Phi}_0)=p(\mathbf{x}|\mathbf{a};\mathbf{\Phi}_0)p(\mathbf{a})$.

To apply Theorem 1 and Corollary 3 to this estimation problem, we have to evaluate the marginal pdf of the data $\mathbf{x}$:

$$p(\mathbf{x};\mathbf{\Phi}_0)=\int p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_0)\,d\mathbf{a}. \tag{3}$$

Unfortunately, in many practical applications, the closed form solution of integral in eq. (3) is extremely difficult or impossible to calculate and this motivates the search for an alternative identifiability criterion. When the marginal pdf of the data $p(\mathbf{x};\mathbf{\Phi}_0)$ is unavailable, we can use the joint pdf $p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_0)$ to define a new alternative identifiability criterion. To this purpose, Definition 2 can be modified as follows:

Definition 5: *Two parameter vectors* $\mathbf{\Phi}_0$ *and* $\mathbf{\Phi}_1$ *(relative to two structures* T_0 *and* T_1*) are said to be observationally equivalent if* $p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_0)=p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_1)$ *for all* $\mathbf{x}\in\mathbb{R}^n$ *and for all* $\mathbf{a}\in\mathbb{R}^l$. $\mathbf{\Phi}_0$ *is otherwise said to be identifiable if there is no other* $\mathbf{\Phi}$ *in* Ω *which is observationally equivalent.*

According to these two definitions, a parameter vector $\mathbf{\Phi}_0$ is non-identifiable if at least another $\mathbf{\Phi}_1$ exists such that:

i. Definition 2: $p(\mathbf{x};\mathbf{\Phi}_0)=p(\mathbf{x};\mathbf{\Phi}_1)$ $\forall\mathbf{x}\in\mathbb{R}^n$ where $p(\mathbf{x};\mathbf{\Phi}_i)=\int p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_i)d\mathbf{a}$, for i=0,1.

ii. Definition 5: $p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_0)=p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_1)$ $\forall\mathbf{x}\in\mathbb{R}^n$, $\forall\mathbf{a}\in\mathbb{R}^l$.

Roughly speaking, Definition 5 requires that the parameter vector $\mathbf{\Phi}$ is identifiable for any realization of $\mathbf{a}$. In Definition 2, $\mathbf{x}$ is observed and $\mathbf{a}$ is averaged out in the pdf, so we do not require it to be observed (known). In the following we derive an operative procedure to verify if, in presence of random nuisance parameters, a parameter vector $\mathbf{\Phi}_0$ is identifiable or not under Definitions 5. Moreover, the relationship between the two definitions of identifiability is investigated.

3.1. Identifiability condition under Definition 5

The aim of this section is to provide a condition to verify if the parameter vector $\mathbf{\Phi}_0$ is identifiable under Definition 5. We want to prove that there is no other parameter vector $\mathbf{\Phi}\in\Omega$, or at least in an open neighborhood of $\mathbf{\Phi}_0$ (local identifiability), such that $p(\mathbf{x},\mathbf{a};\mathbf{\Phi})=p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_0)$. First of all, we have to generalize the definition of the KL divergence under the Definition 5. This can be immediately done by defining a scalar function of $\mathbf{\Phi}$, $H_M(\mathbf{\Phi};\mathbf{\Phi}_0)$, as:

$$\begin{aligned}H_M(\mathbf{\Phi};\mathbf{\Phi}_0)&\triangleq E_{\mathbf{x},\mathbf{a}}\left\{\ln\left(\frac{p(\mathbf{x},\mathbf{a};\mathbf{\Phi})}{p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_0)}\right)\right\}\\&=\int\ln\left(\frac{p(\mathbf{x},\mathbf{a};\mathbf{\Phi})}{p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_0)}\right)p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_0)\,d\mathbf{x}d\mathbf{a}.\end{aligned} \tag{4}$$

Now, we have to show that Theorem 1 holds true under Definition 5 with the generalized definition of KL divergence given in eq. (4). Under Definition 5, Theorem 1 can be recast as follows:

Theorem 2: *Let* $p(\mathbf{x},\mathbf{a};\mathbf{\Phi})$ *and* $p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_0)$ *be two parametric pdfs where* $\mathbf{a}$ *is the vector of the random nuisance parameters. If* $p(\mathbf{x},\mathbf{a};\mathbf{\Phi})=p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_0)$ *for all* $\mathbf{x}\in\mathbb{R}^n$ *and for all* $\mathbf{a}\in\mathbb{R}^l$*, then* $H_M(\mathbf{\Phi};\mathbf{\Phi}_0)=0$. *Otherwise, if* $H_M(\mathbf{\Phi};\mathbf{\Phi}_0)$ *is finite,* $H_M(\mathbf{\Phi};\mathbf{\Phi}_0)<0$.

Proof: The Theorem 2 can be easily proved using the Jensen inequality:

$$\begin{aligned}H_M(\mathbf{\Phi};\mathbf{\Phi}_0)&=E_{\mathbf{x},\mathbf{a}}\left\{\ln\left(\frac{p(\mathbf{x},\mathbf{a};\mathbf{\Phi})}{p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_0)}\right)\right\}\\&\le\ln\left(E_{\mathbf{x},\mathbf{a}}\left\{\frac{p(\mathbf{x},\mathbf{a};\mathbf{\Phi})}{p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_0)}\right\}\right)\\&=\ln\left(\int\frac{p(\mathbf{x},\mathbf{a};\mathbf{\Phi})}{p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_0)}p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_0)\,d\mathbf{x}d\mathbf{a}\right)=\ln(1)=0,\end{aligned} \tag{5}$$

where the equality sign holds if $p(\mathbf{x},\mathbf{a};\mathbf{\Phi})=p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_0)$.

At this point, following the same procedure used in Section 2.2, it is possible to assert that $\mathbf{\Phi}_0$ is globally, or at least locally, identifiable if and only if it is a global, or at least a local, maximum for the KL divergence $H_M(\mathbf{\Phi};\mathbf{\Phi}_0)$. Then we have to show that the gradient of $H_M(\mathbf{\Phi};\mathbf{\Phi}_0)$, evaluated at $\mathbf{\Phi}_0$, i.e. $\nabla(H_M)(\mathbf{\Phi}_0)$, is equal to zero and that the Hessian matrix, also evaluated at $\mathbf{\Phi}_0$, i.e. $[\mathbf{H}(H_M)](\mathbf{\Phi}_0)$, is a negative definite matrix. As proved in [11], it can be shown that the gradient is actually zero and that the Hessian matrix can be expressed as:

$$\begin{aligned}\left[\mathbf{H}(H_M)(\mathbf{\Phi}_0)\right]_{ij}&=\frac{\partial^2}{\partial\Phi_i\partial\Phi_j}H(\mathbf{\Phi};\mathbf{\Phi}_0)\Big|_{\mathbf{\Phi}=\mathbf{\Phi}_0}\\&=-E_{\mathbf{x},\mathbf{a}}\left\{\frac{\partial}{\partial\Phi_i}\ln p(\mathbf{x}|\mathbf{a};\mathbf{\Phi})\Big|_{\mathbf{\Phi}=\mathbf{\Phi}_0}\frac{\partial}{\partial\Phi_j}\ln p(\mathbf{x}|\mathbf{a};\mathbf{\Phi})\Big|_{\mathbf{\Phi}=\mathbf{\Phi}_0}\right\}\\&=E_{\mathbf{x},\mathbf{a}}\left\{\frac{\partial^2}{\partial\Phi_i\partial\Phi_j}\ln p(\mathbf{x}|\mathbf{a};\mathbf{\Phi})\Big|_{\mathbf{\Phi}=\mathbf{\Phi}_0}\right\}=-\left[\mathbf{I}_M(\mathbf{\Phi}_0)\right]_{ij},\end{aligned} \tag{6}$$

where $\mathbf{I}_M(\mathbf{\Phi}_0)$ is the so-called Modified Fisher Information Matrix (MFIM) [6], [7]. Starting from eq. (6), Corollary 3 can be generalized to take the random nuisance parameters into account.

Corollary 4: *Let* $p(\mathbf{x},\mathbf{a};\mathbf{\Phi}_0)$ *be a parametric pdf where* $\mathbf{x}$ *is the data vector and* $\mathbf{a}$ *is the vector of nuisance parameters and let* $\mathbf{I}_M(\mathbf{\Phi}_0)$ *be the MFIM. Then, the parameter vector* $\mathbf{\Phi}_0$ *is locally identifiable if and only if* $\mathbf{I}_M(\mathbf{\Phi}_0)$ *is a full rank matrix.*

3.2. Relationship among the identifiability conditions in presence of random nuisance parameters

The results obtained in the previous sections can be summarized in the following theorem:

Theorem 3: *Let* $p(\mathbf{x},\mathbf{a};\boldsymbol{\Phi}_0)$ *be a parametric pdf where* $\mathbf{x}$ *is the data random vector and* $\mathbf{a}$ *is the random vector of the nuisance parameters, let* $\mathcal{I}_2,\mathcal{I}_5 \subseteq \mathbb{R}^m$ *be the sets of the parameter vectors globally observationally equivalent under Definitions 2 and 5, respectively, and let* $\mathcal{O}_2,\mathcal{O}_5 \subseteq \mathbb{R}^m$ *be two open neighborhoods of* $\boldsymbol{\Phi}_0$ *that contain the parameter vectors locally observationally equivalent to* $\boldsymbol{\Phi}_0$ *under Definitions 2 and 5, respectively. Then, the following relations hold:*

$$\mathcal{I}_5 \subseteq \mathcal{I}_2, \tag{7}$$

$$\mathcal{O}_5 \subseteq \mathcal{O}_2. \tag{8}$$

The proof can be found in [11]. Theorem 3 states that the more restrictive identifiability condition is Definition 2, as expected. This means that, if we use Definition 5 to test the identifiability of a deterministic parameter vector, it might be possible that we classify as identifiable a parameter that it is not identifiable according to Definition 2. However, in a lot of practical estimation problems that involve random nuisance parameters, it is impossible to apply Definition 2 due to the analytical difficulties in the evaluation of the integral in eq. (3). In all these cases, when the classical FIM is impossible to obtain but the MFIM it is easy to evaluate, we can apply Definition 5. Finally, by means of Theorem 3, it is possible to assert that if a parameter vector is not identifiable under Definition 5, then it is not identifiable under Definition 2 as well.

4. IDENTIFIABILITY IN THE RELATIVE GRID-LOCKING PROBLEM FOR NETTED RADAR

The aim of this section is to investigate the identifiability problem in the relative grid-locking problem [8], [9]. The grid-locking problem arises when a set of data coming from two or more sensors must be combined. This problem involves the coordinate transformation and the reciprocal alignment among the various sensors: streams of data from different sensors must be converted into a common coordinate system (or frame) and aligned before they could be used in a tracking or surveillance system. If not corrected, the registration errors can seriously degrade the global system performance by increasing tracking errors and even introducing ghost tracks. A first basic distinction is usually made between relative grid-locking and absolute grid-locking [9]. Here we focus on the relative grid-locking problem, but the application of our findings to the absolute grid-locking problem is also straightforward. First of all, we give some basic concepts on the grid-locking (or sensor registration) problem. One source of registration errors is represented by the sensor calibration (or offset) errors, also called measurement errors. Although the sensors are usually initially calibrated, the calibration may deteriorate over time. There are three measurement errors, one for each component of the measurement vector, i.e. range, azimuth, and elevation. Other kind of registration errors are represented by the attitude (or orientation) errors. Attitude errors can be caused by biases in the gyros of the inertial measurement unit (IMU) of the sensor. There are three possible attitude errors, one for each body-fixed rotation axis. The last source of registration errors is represented by the location (or position) errors caused by bias errors in the navigation system associated with the sensor. In the rest of the paper, we use the following notations:

- *Attitude* biases: we denote by $\boldsymbol{\Theta}_t=(\chi_t\ \psi_t\ \xi_t)^T$, $\boldsymbol{\Theta}_m=(\chi_m\ \psi_m\ \xi_m)^T$ and $d\boldsymbol{\Theta}=(d\chi\ d\psi\ d\xi)^T$ the true attitude angles, the measured attitude angles and the attitude bias errors.
- *Measurement biases*: we denote by $\mathbf{v}_t^k=(\rho_t^k\ \theta_t^k\ \varepsilon_t^k)^T$, $\mathbf{v}_m^k=(\rho_m^k\ \theta_m^k\ \varepsilon_m^k)^T$ and $d\mathbf{v}=(d\rho\ d\theta\ d\varepsilon)^T$ the true target position vector in spherical coordinates, the measured target position vector and the measurement bias errors.
- *Location biases*: we denote by $\mathbf{t}_t=(t_{x,t}\ t_{y,t}\ t_{z,t})^T$, $\mathbf{t}_m=(t_{x,m}\ t_{y,m}\ t_{z,m})^T$ and $d\mathbf{t}=(dt_x\ dt_y\ dt_z)^T$ the true and the measured relative position and the location bias errors.

The convention usually adopted is that the biases must be added to the measured value to obtain the true value of the specific parameter. The bias errors introduced previously can be collected in a vector as:

$$\boldsymbol{\Phi}=\begin{pmatrix}d\rho & d\theta & d\varepsilon & d\chi & d\psi & d\xi & dt_x & dt_y & dt_z\end{pmatrix}^T \tag{9}$$

In the following, we provide an overview of the measurement model. More details can be found in [9]. In the following, the spherical-to-Cartesian transformation is denoted by $\mathbf{h}(\cdot)$ and its inverse, i.e. the Cartesian-to-spherical transformation, by $\mathbf{h}^{-1}(\cdot)$. Both the measurement models of radar #1 and radar #2 involve target position vector $\mathbf{r}_k$ that is directly unobservable and then represents the random nuisance parameter vector. In literature, different models can be found to describe the target position vector, however, the particular model chosen for the target position vector does not affect the identifiability of the error bias vector $\boldsymbol{\Phi}$ to be estimated. Radar #1 is assumed to be ideal, i.e. without bias errors, then its reference system can be taken as the absolute reference system. Under this assumption, the measurement model of radar #1 is given by:

$$\mathbf{v}_{1,m}^k=\mathbf{h}^{-1}(\mathbf{r}_k)+\mathbf{n}_1^k \tag{10}$$

where the noise vector $\mathbf{n}_1^k$ is zero-mean, Gaussian distributed with covariance matrix given by $\mathbf{C}_1=\mathrm{diag}(\sigma_{\rho,1}^2,\ \sigma_{\theta,1}^2,\ \sigma_{\varepsilon,1}^2)$. The measurement model of radar #2 has been derived in [9] as:

$$\begin{aligned}\mathbf{v}_{2,m}^k&=\mathbf{h}^{-1}\left(\mathbf{R}^T\left(\boldsymbol{\Theta}_m+d\boldsymbol{\Theta}\right)\left[\mathbf{r}_k-\left(\mathbf{t}_m+d\mathbf{t}\right)\right]\right)-d\mathbf{v}+\mathbf{n}_2^k\\&\triangleq\boldsymbol{\mu}\left(\mathbf{r}_k,\boldsymbol{\Phi}\right)+\mathbf{n}_2^k,\end{aligned} \tag{11}$$

where $\mathbf{R}$ is a rotation matrix and ^{T} defines the transpose operator. The noise vector $\mathbf{n}_2^k$ is assumed to be zero-mean, Gaussian distributed, independent of $\mathbf{n}_1^k$, with covariance matrix given by $\mathbf{C}_2=\mathrm{diag}(\sigma_{\rho,2}^2,\ \sigma_{\theta,2}^2,\ \sigma_{\varepsilon,2}^2)$. First, we define three sets of K elements as follows:

$$V_1=\left\{\mathbf{v}_{1,m}^k\right\}_{k=1}^K,\ V_2=\left\{\mathbf{v}_{2,m}^k\right\}_{k=1}^K,\ R=\left\{\mathbf{r}_k\right\}_{k=1}^K, \tag{12}$$

where V_1 and V_2 are the sets of the K independent observations coming from radars #1 and #2, and R is the set of the K target positions, which are the random nuisance parameters for our parameter estimation problem. In order to investigate the identifiability of the error bias vector in eq. (9), and then, using Definition 3, to apply Theorem 1 and Corollary 3 to this estimation problem, we have to evaluate the following pdf:

$$p\left(V_1, V_2; \mathbf{\Phi}\right) = \int p\left(V_1, V_2, R; \mathbf{\Phi}\right) dR \qquad (13)$$

As discussed in [9], it is not possible to evaluate in closed form the marginal pdf $p(V_1,V_2;\mathbf{\Phi})$ and only the analytical expression of the joint pdf $p(V_1,V_2,R;\mathbf{\Phi})$ is known. For this reason, the classical definition of identifiability cannot be applied. However, since the joint pdf is known, the MFIM can be easily evaluated, then we adopt the definition of identifiability stated in Definition 5. Following Corollary 4, to investigate the local identifiability under Definition 5, we have to check if the MFIM $\mathbf{I}_M(\mathbf{\Phi})$ is a full rank matrix. First, we start to evaluate the MFIM as:

$$\left[\mathbf{I}_M\left(\mathbf{\Phi}\right)\right]_{ij} = -E_{V_1,V_2,R}\left\{\frac{\partial^2}{\partial \Phi_i \partial \Phi_j} \ln p\left(V_1, V_2 \middle| R; \mathbf{\Phi}\right)\right\}$$
$$= -\sum_{k=1}^{K} E_{\mathbf{v}_{1,m}^k, \mathbf{v}_{2,m}^k, \mathbf{r}_k}\left\{\frac{\partial^2}{\partial \Phi_i \partial \Phi_j} \ln p\left(\mathbf{v}_{1,m}^k, \mathbf{v}_{2,m}^k \middle| \mathbf{r}_k; \mathbf{\Phi}\right)\right\}. \qquad (14)$$

From the measurement models in eqs. (10) and (11), we have that:

$$p\left(\mathbf{v}_{1,m}^k, \mathbf{v}_{2,m}^k \middle| \mathbf{r}_k; \mathbf{\Phi}\right) = p\left(\mathbf{v}_{1,m}^k \middle| \mathbf{r}_k\right) p\left(\mathbf{v}_{2,m}^k \middle| \mathbf{r}_k; \mathbf{\Phi}\right) \qquad (15)$$

where we made explicit the fact that the pdf of $\mathbf{v}^k_{1,m}|\mathbf{r}_k$ does not depend on the parameters vector $\mathbf{\Phi}$. By inserting eq. (15) in eq. (14), we get:

$$\left[\mathbf{I}_M\left(\mathbf{\Phi}\right)\right]_{ij} = \sum_{k=1}^{K} E_{\mathbf{r}_k}\left\{\left[\mathbf{G}\left(\mathbf{r}_k, \mathbf{\Phi}\right)\right]_{ij}\right\}, \qquad (16)$$

where

$$\left[\mathbf{G}\left(\mathbf{r}_k, \mathbf{\Phi}\right)\right]_{ij} \triangleq -E_{\mathbf{v}_{2,m}^k|\mathbf{r}_k}\left\{\frac{\partial^2}{\partial \Phi_i \partial \Phi_j} \ln p\left(\mathbf{v}_{2,m}^k \middle| \mathbf{r}_k; \mathbf{\Phi}\right)\right\} \qquad (17)$$

and, from eq. (11), $\mathbf{v}_{2,m}^k \middle| \mathbf{r}_k \sim \mathcal{N}(\boldsymbol{\mu}(\mathbf{r}_k, \mathbf{\Phi}), \mathbf{C}_2)$. Finally, the entries of matrix $\mathbf{G}(\mathbf{r}_k,\mathbf{\Phi})$ can be evaluated as shown in Appendix 3C of [5]. Unfortunately, the mean value w.r.t. $\mathbf{r}_k$ in eq. (16) cannot be evaluated in closed form due to the analytical complexity of the entries of the matrix $\mathbf{G}(\mathbf{r}_k,\mathbf{\Phi})$, and it is evaluated by running independent Monte Carlo (MC) trials. In our simulations, we used 100 MC trials. Finally, to investigate the local identifiability of the error bias vector $\mathbf{\Phi}$, following Corollary 4, we have to calculate the rank $r_M(\mathbf{\Phi})$ of the MFIM. By using the Matlab® function `rank` to evaluate $r_M(\mathbf{\Phi})$, it can be shown that $r_M(\mathbf{\Phi})=d_{\mathbf{\Phi}}-1$, where $d_{\mathbf{\Phi}}$ is the dimension of the bias vector $\mathbf{\Phi}$. In particular $r_M(\mathbf{\Phi})=8$, while $d_{\mathbf{\Phi}}=9$ (see eq. (9)). This means that the error bias vector $\mathbf{\Phi}$ is not identifiable under Definition 5 and then, from Theorem 3, neither under Definition 2. Such non identifiability can be deduced by the geometry of the relative grid-locking problem [9]. It must be noted in fact that the azimuth measurement bias $d\theta$ and the yaw attitude bias $d\xi$ cannot be distinguished and they should be merged into a single bias. Because of this geometrical coupling, we can define a single bias error as $d\zeta=d\theta+d\xi$. Then we can define a new parameter vector as:

$$\mathbf{\Phi'} = \left(d\rho \quad d\varepsilon \quad d\chi \quad d\psi \quad d\zeta \quad dt_x \quad dt_y \quad dt_z\right)^T. \qquad (18)$$

It can be shown using exactly the same procedure as before that $\mathbf{\Phi'}$ is locally identifiable under Definition 5.

5. CONCLUSIONS

In this paper, we have first discussed and summarized the main concepts on the model identifiability problem. Then, we have generalized the fundamental results on the identifiability to the case where random nuisance parameters are present in the problem. In particular, an alternative definition of identifiability is provided. Such new definition is always applicable, but it is weaker than the classical one. The link between identifiability property and the rank of the classical and modified FIM is investigated. Finally, we applied this new definition of identifiability to the relative grid-locking problem for netted radar system. Future works will explore the possibility to extend these results to the hybrid estimation problem [7] and then to investigate the link between the identifiability and the hybrid FIM.

6. REFERENCES

[1] T. J. Rothenberg, "Identification in parametric models," *Econometrica*, vol. 39, no. 3, pp. 577-591, May 1971.

[2] R. Bowden, "The theory of parametric identification," *Econometrica*, vol. 41, no. 6, pp. 1069-1074, November 1973.

[3] T. C. Koopmans and O. Reiersol, "The Identification of Structural Characteristics," *The Annals of Mathematical Statistics*, vol. 21, no. 2, pp. 165-181, June 1950.

[4] T. Söderström and Stoica P., *System Identification*. London, UK: Prentice-Hall, 1989.

[5] M. S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[6] F. Gini, R. Reggiannini, and U. Mengali, "The Modified Cramér-Rao bound in vector parameter estimation," *IEEE Trans. Commun.*, vol. 46, no. 1, pp. 52-60, January 1998.

[7] Y. Noam, H. Messer, "Notes on the Tightness of the Hybrid Cramér–Rao Lower Bound", *IEEE Trans. Signal Proc.*, vol. 57, no. 6, pp. 2074–2084, June 2009.

[8] M. P. Dana, *Multitarget-Multisensor Tracking: Advanced Applications*. MA: Artech House, 1990.

[9] S. Fortunati, A. Farina, F. Gini, A. Graziano, M. S. Greco and S. Giompapa, "Least Squares estimation and Cramér-Rao Type lower bounds for relative sensor registration process," *IEEE Trans. Signal Proc*., vol. 59, no. 3, pp. 1075-1087, March 2011.

[10] S. Kullback, *Information Theory and Statistics*. New York: John Wiley, 1959.

[11] S. Fortunati, F. Gini, M. S. Greco, A. Farina, A. Graziano, and S. Giompapa, " On the identifiability problem in the presence of random nuisance parameters", *Signal Processing* 92 (2012), pp. 2545-2551.