Semiparametric Inference and Lower Bounds for Real Elliptically Symmetric Distributions

Stefano Fortunati, Member, IEEE, Fulvio Gini, Fellow, IEEE, Maria Sabrina Greco, Fellow, IEEE, Abdelhak M. Zoubir, Fellow, IEEE, and Muralidhar Rangaswamy, Fellow, IEEE

Abstract—This paper has a twofold goal. The first aim is to provide a deeper understanding of the family of the real elliptically symmetric (RES) distributions by investigating their intrinsic semiparametric nature. The second aim is to derive a semiparametric lower bound for the estimation of the parametric component of the model. The RES distributions represent a semiparametric model, where the parametric part is given by the mean vector and by the scatter matrix, while the non-parametric, infinite-dimensional, part is represented by the density generator. Since, in practical applications, we are often interested only in the estimation of the parametric component, the density generator can be considered as nuisance. The first part of this paper is dedicated to conveniently place the RES distributions in the framework of the semiparametric group models. In the second part, building on the mathematical tools previously introduced, the constrained semiparametric Cramér–Rao bound (CSCRB) for the estimation of the mean vector and of the constrained scatter matrix of a RES distributed random vector is introduced. The CSCRB provides a lower bound on the mean squared error of any robust M-estimator of mean vector and scatter matrix when no a priori information on the density generator is available. A closed-form expression for the CSCRB is derived. Finally, in simulations, we assess the statistical efficiency of the Tyler’s and Huber’s scatter matrix M-estimators with respect to the CSCRB.

Index Terms—Parametric model, semiparametric model, Cramér-Rao Bound, semiparametric Cramér-Rao bound, elliptically symmetric distributions, scatter matrix estimation, robust estimation.

I. INTRODUCTION

A PREREQUISITE for any statistical inference method is the notion of a statistical model, say $P$, i.e., a collection, or a family, of probability density functions (pdfs) that is able to characterize random phenomena based on their observations. The most widely used models are parametric models. A parametric model is a family of pdfs parametrized by the elements of a subset $\Gamma$ of a finite-dimensional Euclidean space $\mathbb{R}^q$. The popularity of parametric models is due to the ease of derivation of inference algorithms. On the other hand, a major drawback is their “narrowness” that can lead to misspecification problems [1], [2]. Their counterpart are nonparametric models, a wide family of pdfs that can be required to satisfy some functional constraints, e.g., symmetry, smoothness or moment constraints. While the use of a nonparametric model minimizes the risk of model misspecification, the amount of data needed for nonparametric inference may represent an insurmountable obstacle in practical applications. Semiparametric models have been introduced as a compromise between the “narrowness” of parametric models and the cost of using nonparametric ones [3]. More formally, let $x$ be a random vector taking values in the sample space $\mathcal{X} \subseteq \mathbb{R}^N$. Then, a semiparametric model $\mathcal{P}_{\gamma, l}$ is a family of pdfs parametrized by a finite-dimensional parameter vector of interest $\gamma \in \Gamma \subseteq \mathbb{R}^q$, along with an infinite-dimensional nuisance parameter $l \in \mathcal{L}$, where $\mathcal{L}$ is a set of functions:

$$\mathcal{P}_{\gamma, l} \equiv \{ p_X(x|\gamma, l) : \gamma \in \Gamma, l \in \mathcal{L} \}. \quad (1)$$

There is a rich statistical literature on semiparametric models and their applications. For a comprehensive and detailed list of the main contributions in this field, we refer the reader to [3] and [4] and to the seminal book [5]. However, this profound theoretical understanding of semiparametric models has not been fully exploited in Signal Processing (SP) problems as yet. Two, among the very few, examples of SP applications of the semiparametric inference are the references [6] and [7], where the semiparametric theory has been applied to blind source separation and nonlinear regression, respectively.

This paper aims at improving the understanding of potential applications of semiparametric models. Specifically, we focus our attention on the joint estimation of the mean vector $\mu$ and of the (constrained) scatter matrix $\Sigma$ in the family of Real Elliptically Symmetric (RES) distributions by providing a closed form expression, up to a (numerically performed) singular value decomposition, for the Constrained Semiparametric Cramér–Rao Bound (CSCRB) on the MSE of any robust estimator of $\mu$ and $\Sigma$. As we will discuss below, a constraint on $\Sigma$ is required to avoid the scale ambiguity that characterizes the definition of scatter matrix in RES distributions. The RES class represents a wide family of distributions that includes the Gaussian, the $t$,
the Generalized Gaussian and all the real Compound-Gaussian distributions as special cases ([8]–[14], and [15, Ch. 4]). The elliptical distributions are of fundamental importance in many practical applications since they can be successfully exploited to statistically characterize the non-Gaussian behavior of noisy data. Moreover, as we will discuss below, RES distributions represent an example of a semiparametric model where the parametric part is represented by the mean vector and by the scatter matrix. They should be estimated in the presence of an infinite-dimensional nuisance parameter, i.e., the density generator, which is generally unknown.

This paper is the natural follow on of our previous work [16]. In [16], we provided, in a tutorial and accessible manner, a general introduction to semiparametric inference framework and to the underlying mathematical tools needed for its development. Moreover, the extension of the classical CRB in the presence of a finite-dimensional nuisance deterministic vector to the SCRB where the nuisance parameter belongs to a certain, infinite-dimensional, function space has been reported in [16, Th. 1]. As discussed in the statistical literature and summarized in [16], this generalization can be carried out by means of three key elements:

- the Hilbert space \( \mathcal{H}^q \) of all the \( q \)-dimensional, zero-mean, vector-valued function of the data vector,
- a notion of tangent space \( T \) for a statistical model,
- an orthogonal projection operator on \( T \), i.e., \( \Pi(\cdot|T) \).

A formal definition of the Hilbert space \( \mathcal{H}^q \) and of the projection operator \( \Pi(\cdot|T) \) can be found in Appendix A, while the tangent space for both parametric and semiparametric models has been defined, in a tutorial manner, in [16].

This paper aims at investigating the applications of the general concepts introduced in [16] to the semiparametric model generated by RES distributions. We start by introducing the semiparametric group model and then we continue by showing that the RES class actually possesses this structure. Building on the mathematical framework that characterizes semiparametric group models and, in particular, their tangent space and projection operator, we then show how to derive the CSCRB for the joint estimation of the mean vector \( \mu \) and the scatter matrix \( \Sigma \) of a RES distributed random vector.

The problem of establishing a semiparametric lower bound for the joint estimation of \( \mu \) and \( \Sigma \) in the RES class has been investigated firstly by Bickel in [17], where a bound on the estimation error of the inverse of the scatter matrix has been derived. More discussions and analyses have also been presented in [5, Sec. 4.2 and Sec. 6.3]. More recently, in a series of papers ([18]–[21]), Hallin and Paindavaine rediscovered the RES class as a semiparametric model and, by using two approaches based on Le Cam’s theory [22] and on rank-based invariance [23], they presented the SCRB for the joint estimation of \( \mu \) and \( \Sigma \) in its most general form. However, even if valuable and profound, Hallin and Paindavaine’s work requires a deep understanding of Le Cam’s theory on Local Asymptotic Normality [22]. For this reason, starting from the results in [5, Sec. 4.2 and Sec. 6.3], we propose here an alternative derivation of the SCRB by using a simpler, even if less general, approach.

Along with the derivation of a lower bound on the estimation performance, we always have to specify the class of estimators to which such bound applies. It can be shown that the SCRB is a lower bound to the MSE of any regular and asymptotic linear (RAL) estimator (see [5, Sec. 2.2], [17], [24]–[26], [27, Ch. 3] and [28, Ch. 4] for additional details). Even if we do not address this issue here, it must be highlighted that the class of RAL estimators is a very wide family that encompasses the Maximum Likelihood estimator and all the \( R, S \)-, and in particular, \( M \)-robust estimators.

The rest of the paper is organized as follows. In Section II the semiparametric group model is presented with a particular focus on the calculation of the tangent space and projection operator. Section III collects the basic notions on RES distributions and their intrinsic semiparametric-group structure is investigated. The step-by-step derivation of the SCRB for the estimation of \( \mu \) and \( \Sigma \) is provided in Section IV. The efficiency of the Sample Covariance Matrix and of two robust scatter matrix \( M \)-estimators, Tyler’s and the Huber’s estimators, is assessed in Section V using the previously derived CSCRB. Finally, some concluding remarks are collected in Section VI.

**Notation:** Throughout this paper, italics indicates scalars or scalar-valued functions \((a, A)\), lower case and upper case boldface indicate column vectors \((a)\) and matrices \((A)\) respectively. Note that, since we deal with Hilbert spaces, the word “vector” indicates both Euclidean vectors and vector-valued functions. For clarity, we indicate sometimes a vector-valued function as \( \alpha \equiv \alpha(x) \). Each entry of a matrix \( A \) is indicated as \( a_{i,j} \equiv [A]_{i,j} \). The superscript \( T \) indicates the transpose operator. Finally, for random variables or vectors, the notation \( =_d \) stands for “has the same distribution as”.

II. **The Semiparametric Group Models**

This section introduces a particular semiparametric model, i.e., the semiparametric group model. As the name suggests, this class of semiparametric models is generated by the action of a group of invertible transformations on a random vector whose pdf is allowed to vary in a given set. As we will show in the sequel, this group-based data generating process allows for an easy calculation of the nuisance tangent space and of the orthogonal projection operator. Before introducing the definition of this class of semiparametric models, let us first introduce some related notation.

Let \( A \) be a group of invertible transformations from \( \mathbb{R}^N \) into itself. Suppose that each transformation \( \alpha \in A \) can be parametrized by means of a real vector \( \gamma \in \Gamma \subseteq \mathbb{R}^q \), i.e.

\[
A = \{ \alpha(\cdot; \gamma) \equiv \alpha_\gamma(\cdot); \gamma \in \Gamma \}. \tag{2}
\]

We will indicate with \( \alpha_\gamma^{-1} \) the inverse of \( \alpha_\gamma \). The operation \( \alpha_\gamma \circ \alpha_{\gamma_i} \) denotes the composition of \( \alpha_{\gamma_i} \) and \( \alpha_\gamma \) that can be explicitly expressed as \( (\alpha_\gamma \circ \alpha_{\gamma_i})(\cdot) \equiv \alpha_{\gamma_i}(\alpha_\gamma(\cdot)) \). Finally, \( \gamma_i \) indicates the parameter vector that characterizes the identity transformation \( \alpha_{\gamma_i} \), such that \( \alpha_{\gamma_i}(\cdot) = \cdot \).

**Definition II.1.** (see [5, Sec. 4.2]): Let \( z \in \mathbb{R}^N \) be a real-valued random vector with pdf \( p_Z \), i.e., \( z \sim p_Z(z) \). The parametric model \( P_\gamma \), generated by the action of the group of invertible parametric transformations \( A \), given in (2), on the random vector \( x \sim p_X(x|\gamma) \), is the set of parametric pdfs of the transformed random vector \( \alpha_\gamma(z) = x \sim p_X(x|\gamma) \). Specifically, \( P_\gamma \)
can be explicitly expressed as:
\[
\mathcal{P}_\gamma = \{ p_X | p_X(x|\gamma) = |J(\alpha_\gamma^{-1})(x)|p_Z(\alpha_\gamma^{-1}(x)); \gamma \in \Gamma \},
\]
where \(|J(\alpha_\gamma^{-1})(x)|_{i,j} \triangleq \partial(\alpha_\gamma^{-1}(x;\gamma))/\partial x_i|\partial, j| is the Jacobian matrix of the inverse transformation \(\alpha_\gamma^{-1}\) and \(|\cdot|\) defines the absolute value of a determinant of a matrix. The generalization to semiparametric models can be obtained by allowing the pdf \(p_Z\) to vary within a large set of density functions \(\mathcal{L}\). Consequently, a semiparametric group model generated by the parametric group \(A\) in (2) can be expressed as:
\[
\mathcal{P}_{\gamma;P_Z} = \{ p_X | p_X(x|\gamma,p_Z) = |J(\alpha_\gamma^{-1})(x)|p_Z(\alpha_\gamma^{-1}(x)); \gamma \in \Gamma, p_Z \in \mathcal{L} \}.
\]
Using the notation introduced in [16], the actual “semiparametric vector” is indicated as \((\gamma_0^T, p_{Z,0})^T\) and, consequently, the true pdf is given by:
\[
p_{0}(x) \triangleq p_X(x|\gamma_0, p_{Z,0}) = |J(\alpha_{\gamma_0}^{-1})(x)|p_{Z,0}(\alpha_{\gamma_0}^{-1}(x)).
\]
Moreover, from now on, we denote by \(E_0\{\cdot\}\) the expectation operator with respect to the true pdf \(p_0(x)\). Note that the word “true” here indicates the actual pdf, and consequently the actual semiparametric vector, that characterizes the data.

As mentioned before, the most useful feature of a semiparametric group model is the fact that its underlying group structure allows for a convenient derivation of the nuisance tangent space and of the relevant projection operator. The following proposition formalizes this concept.

Let \(\mathcal{P}_{\gamma;P_Z}\) be a semiparametric group model defined in (4). Let \(T_{\gamma;P_Z}(\gamma)\) be the semiparametric nuisance tangent space of \(\mathcal{P}_{\gamma;P_Z}\) evaluated at \((\gamma_0^T, p_{Z,0})^T\), where \(\gamma\) is a generic element of the finite-dimensional parameter space \(\Gamma\). Let us indicate by \(T_{\gamma;P_Z}(\gamma_e)\) the semiparametric nuisance tangent space of \(\mathcal{P}_{\gamma;P_Z}\) evaluated at \((\gamma_e^T, p_{Z,e})^T\), where, as defined before, \(\gamma_e\) is the parameter vector that characterizes the identity transformation.

**Proposition II.1:** Let \(t\) be a generic \(q\)-dimensional vector-valued function belonging to \(T_{\gamma;P_Z}(\gamma_e)\), then the semiparametric nuisance tangent space \(T_{\gamma;P_Z}(\gamma)\) can be obtained from \(T_{\gamma;P_Z}(\gamma_e)\) as follows:
\[
T_{\gamma;P_Z}(\gamma) = \{ t \circ \alpha^{-1}\gamma; t \in T_{\gamma;P_Z}(\gamma_e) \}, \forall \gamma \in \Gamma.
\]

Moreover, let \(I\) be a generic \(q\)-dimensional vector-valued function in \(\mathcal{H}^q\), then the projection operator \(\Pi(\cdot|T_{\gamma;P_Z}(\gamma))\) on \(T_{\gamma;P_Z}(\gamma)\) (see Appendix A) can be obtained from the projection operator \(\Pi(\cdot|T_{\gamma;P_Z}(\gamma_e))\) on \(T_{\gamma;P_Z}(\gamma_e)\) as follows:
\[
\Pi(\cdot|T_{\gamma;P_Z}(\gamma)) = \Pi(\cdot|\alpha^{-1}\gamma|T_{\gamma;P_Z}(\gamma_e)) \circ \alpha^{-1}\gamma, \forall \gamma \in \Gamma.
\]

The proof can be found in [5, Sec. 4.2, Lemma 3].

It is worth noticing that Proposition II.1 can be directly used to derive the nuisance tangent space at the true semiparametric vector \((\gamma_0^T, p_{Z,0})^T\), i.e., \(T_{\gamma;P_Z}(\gamma_0) \equiv T_{\gamma;P_Z}(\gamma_0)\) and the relevant projection operator \(\Pi(\cdot|T_{\gamma;P_Z}(\gamma_0))\). This can be done by evaluating the relations (6) and (7) at the true parameter vector of interest \(\gamma_0\). As discussed below, Proposition II.1 is of fundamental importance for the derivation of the Semiparametric Cramér-Rao Bound (SCRB) for \(\gamma_0\) in the class of RES distributions.

### III. THE FAMILY OF RES DISTRIBUTIONS AS A SEMIPARAMETRIC MODEL

In this section, the semiparametric nature of the family of RES distributions is investigated. In particular, we show that the RES class can be conveniently interpreted as a semiparametric group model. Here we restrict the discussion to the absolutely continuous case [14, Sec. III.D], i.e., we always assume that a RES distributed random vector admits a pdf. In what follows, we exploit the definition of the RES class provided in [9, Ch. 3] and [29], since it is particularly useful for our aims.

#### A. Spherically Symmetric (SS) Distributions

As a prerequisite for the definition of the RES class, Spherically Symmetric (SS) distributions need to be introduced first.

**Definition III.1:** Let \(z \in \mathbb{R}^N\) be a real-valued random vector and let \(O\) be the set of all the orthogonal transformations such that:
\[
O \ni O : \mathbb{R}^N \rightarrow \mathbb{R}^N
\]
\[
z \mapsto O(z) = Oz,
\]
for any orthogonal matrix \(O\), i.e., for any \(O \in \mathbb{R}^{N \times N}\) such that \(O^T O = OO^T = I\). Then, \(z\) is said to be SS-distributed if its distribution is invariant to any orthogonal transformations \(O \in \mathcal{O}\), i.e.,
\[
z =_d Oz.
\]
We indicate with \(\mathcal{S}\) the class of all SS-distributions.

As a consequence, the following properties hold true (see, e.g., [9, Ch. 3] or [29] for the proof):

- **P1** The SS-distributed random vector \(z \sim SS(g)\) has a pdf given by:
\[
p_Z(z) = 2^{-N/2} g \left( ||z||^2 \right),
\]
where \(G \ni g\) is a function, called density generator, that depends on \(z\) only through \(||z||\) and
\[
G = \left\{ g : \mathbb{R}^+ \rightarrow \mathbb{R}^+ \left| \int_0^\infty t^{N/2-1} g(t) dt < \infty \right. \right\}
\]
where the integrability condition in (11) is required to guarantee the integrability of \(p_Z(z)\) (see [9, eq. 3.25]). Consequently, the set of all SS pdfs can be described as:
\[
\mathcal{S} = \{ p_Z | p_Z(z) = 2^{-N/2} g \left( ||z||^2 \right), \forall g \in \mathcal{G} \}.
\]
- **P2** Let \(s_N \triangleq 2\pi^{N/2}/\Gamma(N/2)\) be the surface area of the unit sphere in \(\mathbb{R}^N\), then the pdf of the random variables \(Q =_d ||z||^2\) and \(R \triangleq \sqrt{Q}\), called second-order modular variate and modular variate respectively [14], are given by:
\[
p_Q(q) = s_N 2^{-N/2-1} q^{N/2-1} g \left( q \right),
\]
\[
p_R(r) = s_N 2^{-N/2} r^{N-1} g \left( r^2 \right).
\]
- **P3** The Stochastic Representation Theorem: Let \(u \sim \mathcal{U}(\mathbb{R}^S)\) be a random vector uniformly distributed on the real unit sphere of dimension \(S\), indicated as \(\mathbb{R}^S\).
If \( z \in \mathbb{R}^N \) is SS-distributed with pdf given by (10), then:
\[
z = d \sqrt{u} = d \mathcal{R}u,
\]
(15)
where \( Q \sim p_Q(q) \) in (13), \( \mathcal{R} \sim p_R(r) \) in (14). Moreover, \( Q \) and \( u \) (or \( \mathcal{R} \) and \( u \)) are independent. Note that \( u \) in (15) satisfies the following three properties: \(|u| = 1\), \( E\{u\} = 0 \) and \( E\{uu^T\} = N^{-1}\mathbf{I} \).

P4) Maximal Invariant Statistic. The Stochastic Representation Theorem shows that there exists a one-to-one relationship between every \( z \sim SS(g) \) and every couple \((\mathcal{R}, u)\). Moreover, it is easy to verify that the modular variate \( \mathcal{R} \) is a maximal invariant statistic for the set of the SS-distributed random vectors.  

We can now introduce the class of RES distributions as a semiparametric group model.

B. The RES Class as a Semiparametric Group Model

At first, let us define the parameter space \( \Omega \subseteq \mathbb{R}^q \) of dimension \( q = N(N + 3)/2 \) as:
\[
\Omega = \{ \phi \in \mathbb{R}^q | \phi = (\mu^T, \text{vecs}(\Sigma^T))^T; \mu \in \mathbb{R}^N, \Sigma \in \mathcal{M}_N \},
\]
where \( \mu \in \mathbb{R}^N \) is a real-valued \( N \)-dimensional vector while \( \Sigma \) is an \( N \times N \) matrix belonging to the set \( \mathcal{M}_N \) of all the symmetric, positive-definite matrices of dimension \( N \times N \). Note that the vecs operator maps the \( N \times N \) symmetric matrix \( \Sigma \) in an \( (N + 1)/2 \)-dimensional vector of the entries of the lower triangular sub-matrix of \( \Sigma \) [31], [32]. We can now introduce the group \( A \) of all affine transformations, parameterized by the parameter space \( \Omega \) in (16), such that:
\[
A \ni \alpha_\phi : \mathbb{R}^N \rightarrow \mathbb{R}^N ; \forall \phi \in \Omega
\]
\[
z \mapsto \alpha_\phi(z) = \mu + \Sigma^{1/2}z.
\]
(17)
The identity element \( \alpha_\phi(\cdot) \) of the group \( A \) is parameterized by the vector \( \phi_0 = (0^T, \text{vecs}(\mathbf{I}^T))^T \), while the inverse transformation is given by:
\[
\alpha_\phi^{-1}(\cdot) = \Sigma^{-1/2}(\cdot - \mu).
\]
(18)
The class of RES distributions is defined as the class of distributions that is closed under the action of the group \( A \) in (17) on any SS-distributed random vector. The next definition formalizes this statement.

Definition III.2: A real-valued random vector \( x \in \mathbb{R}^N \) is said to be RES-distributed with mean value \( \mu \) and scatter matrix \( \Sigma \), if it can be expressed as:
\[
x = \alpha_\phi(z) = \mu + \Sigma^{1/2}z = d \mu + \sqrt{\Omega}\Sigma^{1/2}u,
\]
(19)
where \( \phi \in \Omega \) defined in (16), \( z \sim SS(g) \) is an SS-distributed random vector, while \( u \sim \mathcal{U}(\mathbb{R}, S^N) \) and the 2nd-order modular variate \( Q \) have been already defined in (15). In particular:
\[
Q = d \|z\|^2 = \|\alpha_\phi^{-1}(x)\|^2 = (x - \mu)^T\Sigma^{-1}(x - \mu) \equiv Q.
\]
(20)

We refer the reader to [9, Ch. 3], [14], and [29] for the proof. Definition III.2 provides the link between the RES family and the semiparametric group model defined in Section II. As a consequence, the explicit expression of the pdf of an RES distributed random vector can be obtained as shown in the Definition II.1, Equation (4). In particular, the determinant of the Jacobian matrix of the inverse transformation in (18) is \( |J(\alpha_\phi^{-1})| = |\Sigma^{-1/2}| = |\Sigma|^{-1/2} \). Then, the pdf of any RES-distributed random vector \( x \) can be obtained from the relevant SS-distributed random vector \( z \), i.e., \( p_x(z) \) in (10), as:
\[
RES_N(x; \mu, \Sigma, g) = \frac{1}{|\Sigma|^{1/2}p_Z(\alpha_\phi^{-1}(x))}
\]
\[
= 2^{-N/2} |\Sigma|^{-1/2} g((x - \mu)^T\Sigma^{-1}(x - \mu)), \forall g \in \mathcal{G}.
\]
(21)

Moreover, the general description of a semiparametric group model given in (4) can be specialized for the RES case as:
\[
P_{\phi,g} = \{ p_X | p_X(x|\phi, g) = 2^{-N/2} |\Sigma|^{-1/2} g(||\alpha_\phi^{-1}(x)||^2) ; \phi \in \Omega, g \in \mathcal{G} \},
\]
(22)
where \( \mathcal{G} \) is the set of density generators given in (11). Clearly, the mean vector of \( x \sim RES_N(x; \mu, \Sigma, g) \) is given by \( E\{x\} = \mu \) while, if \( E\{Q\} < \infty \), its covariance matrix \( M \equiv M = N^{-1}E\{Q\}\Sigma \).

As extensively discussed in the literature on elliptically symmetric distributions, the representation of an RES distributed vector \( x \) is not uniquely determined by (19). In fact, \( x = d \mu + \sqrt{\Omega}\Sigma^{1/2}u = d\mu + \sqrt{\mathcal{G}}(c\Sigma^{1/2}u), \forall c > 0 \). This scale ambiguity can also be seen as a consequence of the functional form of an RES pdf given in (21) since \( RES_N(x; \mu, \Sigma, g(t)) \equiv RES_N(x; \mu, c^2\Sigma, g(t/c^2)), \forall c > 0 \). To avoid this well-known identifiability problem, we impose the following constraint on the trace of \( \Sigma \), i.e.,
\[
c(\Sigma) = \text{tr}(\Sigma) = N = 0.
\]
(23)
This constraint limits the parameter vector \( \phi \in \Omega \), where \( \Omega \) is defined in (16), in a lower dimensional smooth manifold
\[
\Omega = \{ \phi \in \Omega | \text{tr}(\Sigma) = N \},
\]
(24)
whose dimension is \( d = q - 1 \). Trace constraint is only an example of all the possible constraints that can be imposed on the scatter matrix to avoid scale ambiguity. For a deep and insightful analysis of the impact of the particular constraint on \( \Sigma \) on the estimation performance, we refer to [20] and [21].
The properties of the semiparametric group model previously discussed can be exploited to derive the CSCRB for the estimation of the constrained parameter vector \( \phi = (\mu^T, \text{vecs}(\Sigma)^T)^T \in \Omega \), where \( \Omega \) given in (24).

**IV. The Constrained Semiparametric Cramér-Rao Bound for the RES Class**

This section is devoted to the derivation of a closed form expression of the CSCRB for the estimation of \( \phi \in \Omega \). The theoretical foundation of the generalization of the Cramér-Rao inequality in the semiparametric framework can be found in [5, Sec. 3.4], [24], [27, Th. 4.1], and [33]. Moreover, in [16], it is shown, in a tutorial manner, how the CSCRB can be obtained as a result of a limit process of the classical CRB derived in the presence of a finite-dimensional nuisance parameter vector.

As explained in the previous section, to avoid the scale ambiguity of the RES class, we need to put a constraint on the scatter matrix. In order to take this requirement into account, we propose in the sequel the extension of [16, Th. 1] to the case of constrained, finite-dimensional, parameter vector. In particular, suppose that the finite-dimensional parameter vector of interest \( \gamma_0 \in \Gamma \subseteq \mathbb{R}^q \) is required to satisfy \( k \) (with \( k < q \)) continuously differentiable constraints ([34]–[36]):

\[
c(\gamma_0) = 0.
\]

This set of constraints define a smooth manifold, of dimension \( q - k \), in the parameter space \( \Gamma \), such that:

\[
\hat{\Gamma} = \{ \gamma \in \Gamma \subseteq \mathbb{R}^q | c(\gamma) = 0 \}
\]

Moreover, suppose that the \( k \times q \) Jacobian matrix of the constraints, defined as \( J_c(\gamma)_{i,j} = \partial c_i(\gamma)/\partial \gamma_j \) has full row rank for any \( \gamma \in \Gamma \) satisfying (25). Consequently, there exists a matrix \( U \in \mathbb{R}^{k \times q} \) whose columns form an orthonormal basis for the null space of \( J_c(\gamma_0) \), i.e.,

\[
J_c(\gamma_0) U = 0_{k \times q}, \quad U^T U = I.
\]

**Theorem IV.1:** The Constrained Semiparametric Cramér-Rao Bound (CSCRB) for the estimation of the constrained finite-dimensional vector \( \gamma_0 \in \hat{\Gamma} \) in the presence of the nuisance function \( l_0 \in \mathcal{L} \) is given by:

\[
\text{CSCRB}(\gamma_0 | l_0) = U (U^T \hat{I}(\gamma_0 | l_0) U)^{-1} U^T,
\]

where:

\[
\hat{I}(\gamma_0 | l_0) = \left[ E_0 \{ \bar{s}_0 \bar{s}_0^T \} \right],
\]

is the semiparametric Fisher Information Matrix (SFIM) and \( \bar{s}_0 \) is the semiparametric efficient score vector defined as:

\[
\bar{s}_0 = s_{\gamma_0} - \Pi(s_{\gamma_0} | T_{l_0}),
\]

where \( \Pi(s_{\gamma_0} | T_{l_0}) \) is the orthogonal projection of the score vector of the parameters of interest \( s_{\gamma_0} \) on the semiparametric nuisance tangent space \( T_{l_0} \), evaluated at \( l_0 \). Finally, matrix \( U \) is defined in (27). Note that \( \bar{s}_0, s_{\gamma_0} \) and \( \Pi(s_{\gamma_0} | T_{l_0}) \) are \( q \)-dimensional functions of the observation vector \( x \).

The proof of the “unconstrained” part of this theorem can be found in [24] and in [27, Th. 4.1], while a more abstract and general formulation can be found in [5, Sec. 3.4] and in [33]. The proof of the “constrained” part can be obtained through a straightforward application of the approach discussed in [35] for the constrained CRB.

The remainder of this section is devoted to the evaluation of the CSCRB in (28) for the estimation of the mean vector \( \mu \) and of the constrained scatter matrix \( \Sigma \), such that \( \text{tr}(\Sigma) = N' \), of a RES distributed random vector \( x \sim \text{RES}_N(\mu_0, \Sigma_0, g_0) \) in the presence of the nuisance function \( g_0 \in \mathcal{G} \). To this end, according to Theorem IV.1, we have to evaluate:

a) the score vector of the parameters of interest \( s_{\phi_0} \equiv s_{\phi_0}(x) \), where \( \phi_0 = (\mu_0^T, \text{vecs}(\Sigma_0)^T)^T \in \Omega \) given in (16),

b) the projection operator \( \Pi(s_{\phi_0} | T_{g_0}) \), where \( T_{g_0} \equiv T_{g_0}(\phi_0) \) is the semiparametric nuisance tangent space for the RES distributions class evaluated at the true semiparametric vector \( (\phi_0^T, g_0)^T \),

c) the semiparametric efficient score vector \( s_{\theta_0} \equiv s_{\theta_0}(x) \) in (30) and the SFIM \( \hat{I}(\phi_0 | g_0) \) in (29),

d) the matrix \( U \) in (27) and then the CSCRB in (28).

In what follows, we will provide this calculation step-by-step.

**A. Evaluation of the Score Vector \( s_{\phi_0}(x) \)**

By definition, the score vector of the parameters of interest is given by:

\[
s_{\phi_0}(x) = \nabla_{\phi} \ln p_X(x | \phi_0, g_0) = \begin{pmatrix} s_{\mu_0}(x) \\ s_{\text{vecs}(\Sigma_0)}(x) \end{pmatrix},
\]

where \( \phi_0 \in \Omega \) and \( \Omega \) is the unconstrained parameter space defined in (16) and:

\[
s_{\mu_0}(x) = \nabla_{\mu} \ln p_X(x | \mu_0, \Sigma_0, g_0),
\]

\[
s_{\text{vecs}(\Sigma_0)}(x) = \nabla_{\text{vecs}(\Sigma)} \ln p_X(x | \mu_0, \Sigma_0, g_0),
\]

and where \( \mu_0, \Sigma_0 \) and \( g_0 \) represents the true mean vector, the true scatter matrix and the true density generator, respectively. By substituting in (32) the explicit expression of \( p_X(x | \mu_0, \Sigma_0, g_0) \) given in (21) and by exploiting the differentiation rules with respect to vector and matrices provided e.g., in [37, Ch. 8], we have that:

\[
s_{\mu_0}(x) = -2 \psi_0(Q_0) \Sigma_0^{-1} (x - \mu_0)
\]

\[
= d -2 \sqrt{Q_0}(Q) \Sigma_0^{-1/2} u
\]

where, according to (20),

\[
Q_0 = \left\| \alpha_{\phi_0}^{-1}(x) \right\|^2 = (x - \mu_0)^T \Sigma_0^{-1} (x - \mu_0) = d Q,
\]

and the last two equalities follow directly from the Stochastic Representation of a RES vector given in (19) and

\[
\psi_0(t) \triangleq \frac{1}{g_0(t)} \frac{dg_0(t)}{dt}.
\]

Similarly, the term \( s_{\text{vecs}(\Sigma_0)}(x) \) in (33) can be evaluated by applying the rules of the differential matrix calculus detailed in...
and by using standard properties of the Kronecker product and of the vec operator ([31], [32], [37], Ch. 15]) as:

\[ s_{\text{vec}}(\Sigma_0)(x) = -D_N^T \left( 2^{-1} \text{vec}(\Sigma_0^{-1}) + \psi_0(Q_0)\Sigma_0^{-1} \right) \]

\[ \otimes \Sigma_0^{-1/2} \text{vec}(x - \mu_0) \]

\[ = -D_N^T \left( 2^{-1} \text{vec}(\Sigma_0^{-1}) + \mathcal{Q}\psi_0(Q)\Sigma_0^{-1/2} \right) \]

\[ \otimes \Sigma_0^{-1/2} \text{vec}(uu^T) \]  

(37)

As before, the second equality follows from the Stochastic Representation (19) while \( D_N \) is the \( N^2 \times N(N + 1)/2 \) duplication matrix that is implicitly defined by the equality \( D_N \text{vec}(A) = \text{vec}(A) \) for any \( N^2 \times N^2 \) symmetric matrix \( A \) ([31], [32]). See also [38]–[40] for similar calculation.

**B. Evaluation of the Projection Operator \( \Pi(s_{\phi_0}|T_{g_0}) \)**

In order to obtain the explicit expression of the projection operator, we will exploit the fact that the RES class is a semiparametric group model. In particular, \( \Pi(s_{\phi_0}|T_{g_0}) \) can be obtained by specializing Proposition II.1 for the RES distributions.

For the sake of clarity, let us recall that, according to the definition of the group \( A \) of the affine transformations (17), for any \( \phi_0 \in \Omega \), we have:

\[ \alpha_{\phi_0}(z) \triangleq \mu_0 + \Sigma_0^{1/2} z, \quad z \sim SS(g_0), \]

\[ (38) \]

\[ \alpha_{\phi_0}^{-1}(x) \triangleq \Sigma_0^{1/2}(x - \mu_0), \quad x \sim RES_N(x; \mu_0, \Sigma_0, g_0). \]

\[ (39) \]

Let \( T_{g_0} \equiv T_{g_0}(\phi_0) \) and \( T_{\Sigma_0} \equiv T_{g_0}(\phi_e) \) be the semiparametric nuisance tangent spaces of the RES class evaluated at the true semiparametric vector \((\phi_0^*, g_0^*)^T\) and at \((\phi_e^*, g_0^*)^T\), where \( \phi_e = (0^T, \text{vec}(I)^T)^T \) is the vector that characterizes the identity transformation given in (39). Note that, since it is evaluated at the identity transformation, \( T_{\Sigma_0} \) can be interpreted as the tangent space of the SS distribution class evaluated at the true \( g_0 \), and this explains the chosen notation. Now, by directly applying Proposition II.1, we have that the projection of the score vector of the parameters of interest \( s_{\phi_0}(x) \) on the tangent space \( T_{g_0} \) can be expressed as:

\[ \Pi(s_{\phi_0}(x)|T_{g_0})(x) = \left( \Pi((s_{\phi_0} \circ \alpha_{\phi_0})(z)|T_{\Sigma_0}) \circ \alpha^{-1}_{\phi_0}(x) \right)(x), \]

(41)

where \( x \sim RES_N(x; \mu_0, \Sigma_0, g_0) \) and \( z \sim SS(g_0) \).

In what follows, we show how to obtain an explicit expression of the projection in (41).

1) **Calculation of \( s_{\phi_0} \circ \alpha_{\phi_0}(z) \):** From a visual inspection of the expressions in (34) and (37), one can be rewritten as a function of the inverse transformation \( \alpha^{-1}_{\phi_0}(x) \), i.e.:

\[ s_{\mu_0}(x) = (s_{\mu_0} \circ \alpha^{-1}_{\phi_0})(x) = -2\psi_0(Q_0)\Sigma_0^{-1/2} \alpha^{-1}_{\phi_0}(x), \]

(42)

\[ \text{and} \]

\[ s_{\text{vec}}(\Sigma_0)(x) = (s_{\text{vec}}(\Sigma_0) \circ \alpha^{-1}_{\phi_0})(x) \]

\[ = -D_N^T \left( 2^{-1} \text{vec}(\Sigma_0^{-1}) + \psi_0(Q_0)\Sigma_0^{-1/2} \right) \]

\[ \otimes \Sigma_0^{-1/2} \text{vec}(uu^T) \]  

(43)

Then, we have that:

\[ (s_{\phi_0} \circ \alpha_{\phi_0})(z) \]

\[ = \left( \left( \begin{array}{c} (s_{\mu_0} \circ \alpha^{-1}_{\phi_0}(\alpha_{\phi_0}(z)) \\ (s_{\text{vec}}(\Sigma_0) \circ \alpha^{-1}_{\phi_0}(\alpha_{\phi_0}(z)) \end{array} \right) \right), \]

(44)

and

\[ \tilde{s}_{\mu_0}(z) = -2\psi_0(Q_0)\Sigma_0^{-1/2} z = -2\sqrt{\mathcal{Q}}\psi_0(Q)\Sigma_0^{-1/2} u \]

(45)

\[ \tilde{s}_{\text{vec}}(\Sigma_0)(z) = -D_N^T \left( 2^{-1} \text{vec}(\Sigma_0^{-1}) + \psi_0(Q_0)\Sigma_0^{-1/2} \right) \]

\[ \otimes \Sigma_0^{-1/2} \text{vec}(z z^T) \]

\[ = -D_N^T \left( 2^{-1} \text{vec}(\Sigma_0^{-1}) + \mathcal{Q}\psi_0(Q)\Sigma_0^{-1/2} \right) \]

\[ \otimes \Sigma_0^{-1/2} \text{vec}(uu^T) \]  

(46)

2) **Derivation of \( T_{\Sigma_0} \):** The next step is the evaluation of \( T_{\Sigma_0} \), i.e., the tangent space of \( S \) in (12), evaluated at the true density generator \( g_0 \). Using the procedure discussed in [5, Appendix A.3], it is possible to verify that \( T_{\Sigma_0} \) is a \( q \)-replicating Hilbert space \( T_{\Sigma_0} = T \times \cdots \times T \) (see also Appendix A) such that:

\[ T_{\Sigma_0} \triangleq \{a|a \text{ is any vector in } \mathbb{R}^q, l \in T \}, \]

(47)

where

\[ T = \{l \in H^l | l \text{ is invariant under } O \} = \{l|l(z) = l(|z|), \quad E_0\{l(|z|)\} = 0\}, \]

(48)

and \( O \) is the group of orthogonal transformations defined in (8).

Let us now recall that, according to Property P4 in Section III, the modular variate \( \mathcal{R} \) is a maximal invariant statistic for an SS distribution. Then, by using the procedure discussed in [5, Sec. 6.3, Example 1] and in accordance with the discussion provided in Appendix B, we have that the projection operator \( \Pi(\cdot|T_{\Sigma_0}) \) on the tangent space \( T_{\Sigma_0} \) can be obtained as the expectation operator \( E_{0|R}\{\cdot|R\} \) with respect to the maximal invariant statistic \( R \), i.e.,

\[ \Pi(\cdot|T_{\Sigma_0}) = E_{0|R}\{\cdot|R\}, \quad \forall l \in H^q. \]

(49)

3) **Calculation of the Projection Operator \( \Pi(s_{\phi_0}|T_{g_0}) \):** In order to evaluate \( \Pi(\cdot|T_{g_0}) \), we can use the property of the semiparametric group models given in (41). Let us start by deriving the expression for \( \Pi((s_{\phi_0} \circ \alpha_{\phi_0})|T_{\Sigma_0}) = \Pi(s_{\phi_0}|T_{\Sigma_0}) \). By exploiting the results collected in the previous subsections, \( \Pi(s_{\phi_0}|T_{\Sigma_0}) \) can be easily obtained by substituting (45) and (46) in (49) and then
evaluating the conditional expectation operator. Specifically:
\[
\Pi(\hat{s}_{\mu_0} | T_{S_0}) = E_{0|R}(\hat{s}_{\mu_0} | R) \\
= d - 2\sqrt{Q}\psi_0(Q)\Sigma_0^{-1/2}E\{u\} = 0, \quad (50)
\]
and
\[
\Pi(\hat{s}_{\text{vecs}(\Sigma_0)} | T_{S_0}) = E_{0|R}(\hat{s}_{\text{vecs}(\Sigma_0)} | R) \\
= d - D_N^T \left( 2^{-1} \text{vec}(\Sigma_0^{-1}) + Q\psi_0(Q)\Sigma_0^{-1/2} \otimes \Sigma_0^{-1/2} E\{uu^T\} \right) \\
= - D_N^T \left( \frac{1}{2} + \frac{1}{N} Q\psi_0(Q) \right) \text{vec}(\Sigma_0^{-1}), \quad (51)
\]
then, consequently:
\[
\Pi(\hat{s}_{\phi_0} | T_{S_0}) = d \left( \begin{array}{c}
0 \\
-D_N^T \left( \frac{1}{2} + \frac{1}{N} Q\psi_0(Q) \right) \text{vec}(\Sigma_0^{-1})
\end{array} \right) \\
= d \Pi(\hat{s}_{\phi_0} | T_{g_0}), \quad (52)
\]
where the last equality follows from (41) and from the fact that
\[\Pi(\hat{s}_{\phi_0} | T_{S_0})\] does not depend on \(z\).

Now, a comment is in order. Equation (52) tells us that the score function of the mean value is orthogonal to the nuisance tangent space \(T_{g_0}\). This means that not knowing the true density generator does not have any impact in the (asymptotic) estimation performance of the mean vector \(\mu_0\) [17].

C. Calculation of the Semiparametric Efficient Score Vector \(\hat{s}_0(x)\) and of the SFIM \(\bar{I}(\phi_0, g_0)\)

By collecting the previous results, the semiparametric efficient score vector in (30) can be evaluated as:
\[
\hat{s}_0(x) = s_{\phi_0}(x) - \Pi(s_{\phi_0}(x)|T_{g_0})(x) = \left( \begin{array}{c}
\hat{s}_{\mu_0}(x) \\
\hat{s}_{\text{vecs}(\Sigma_0)}(x)
\end{array} \right) = d \\
\left( \begin{array}{c}
-2\sqrt{Q}\psi_0(Q)\Sigma_0^{-1/2}u \\
-D_N^T Q\psi_0(Q) \left( \Sigma_0^{-1/2} \otimes \Sigma_0^{-1/2} \text{vec}(uu^T) - \frac{1}{N} \text{vec}(\Sigma_0^{-1}) \right)
\end{array} \right), \quad (53)
\]
Finally, through direct calculation, the SFIM \(\bar{I}(\phi_0, g_0)\) can be obtained as:
\[
\bar{I}(\phi_0, g_0) = E_0 \{\hat{s}_0(x)\hat{s}_0(x)^T\} \\
= \left( \begin{array}{cc}
C_0(\hat{s}_{\mu_0}) & 0 \\
0^T & C_0(\hat{s}_{\text{vecs}(\Sigma_0)})
\end{array} \right), \quad (54)
\]
where \(C_0(I) \triangleq E_0 \{II^T\}, \forall I \in H^d\). Note that the off-diagonal block matrices in (54) are nil because all the third-order moments of \(u\) vanish [14, Lemma 1]. From the previous results and using some algebra, we get:
\[
C_0(\hat{s}_{\mu_0}) = \frac{4E\{Q^2\psi_0(Q)^2\}N}{N} \Sigma_0^{-1}, \quad (55)
\]
and
\[
C_0(\hat{s}_{\text{vecs}(\Sigma_0)}) = \frac{2E\{Q^2\psi_0(Q)^2\}}{N(N+2)} D_N^T \left( \Sigma_0^{-1} \otimes \Sigma_0^{-1} - \frac{1}{N} \text{vec}(\Sigma_0^{-1}) \text{vec}(\Sigma_0^{-1})^T \right) D_N. \quad (56)
\]
Note that we had not imposed the constraint on the scatter matrix \(\Sigma_0\) as yet. In particular, in all the previous equations, \(\Sigma_0\) can be considered as the unconstrained scatter matrix. The next subsection is then dedicated to the derivation of the SCRB for the constrained parameter vector.

D. Evaluation of the SCRB(\(\phi_0|g_0\))

As showed in Theorem IV.1, as a prerequisite of the derivation of the SCRB on the estimation of \(\phi_0 \in \Omega\), we have to calculate the matrix \(U\) defined in (27). This can be done by using the same procedure discussed in [41]. Specifically, let us start by evaluating the gradient of the constraint in (23) as:
\[
J_c(\Sigma_0) = \nabla_{\text{vecs}(\Sigma_0)} c(\Sigma_0) = 1^T I, \quad (57)
\]
where \(1_I\) is the \(N(N + 1)/2\)-dimensional column vector defined as:
\[
1_I(i) = \begin{cases} 
1 & i \in I \\
0 & \text{otherwise}
\end{cases}, \quad (58)
\]
where \(I = \{i | i = 1 + N(j - 1) - (j - 1)(j - 2)/2, j = 1, \ldots, N\}\).

Then, \(U\) can be obtained by numerically evaluating, using singular value decomposition (SVD), the \(q = 1\) orthonormal eigenvectors associated with the zero eigenvalue of \(1_I\).

Finally, the constrained SCRB (CSRB) for the estimation of \(\phi_0 \in \Omega\) in (24) can be expressed as:
\[
\text{CSRB}(\phi_0|g_0) = \left( \begin{array}{cc}
\frac{N}{4E\{Q\psi_0(Q)^2\}} \Sigma_0 & 0 \\
0 & U(U^T C_0(\hat{s}_{\text{vecs}(\Sigma_0)})U)^{-1}U^T
\end{array} \right). \quad (59)
\]
Note that the block-diagonal structure of \(\text{CSRB}(\phi_0|g_0)\) implies that not knowing the mean vector \(\mu_0\) does not have any impact on the asymptotic performance in the estimation of the scatter matrix \(\Sigma_0\). From a practical point of view, this also means that the unknown \(\mu_0\) can be substituted with any consistent estimator without affecting the asymptotically optimal performance of the scatter matrix estimator.

V. SIMULATION RESULTS FOR RES DISTRIBUTED DATA

In this section, we investigate the efficiency of three well-known scatter matrix estimators with respect to the SCRB: the constrained Sample Covariance Matrix (CSCM) estimator, the constrained Tyler's (C-Tyler) estimator and the constrained Huber's (C-Hub) estimator. Note that none of these estimators relies on the a-priori knowledge of the true density generator \(g_0 \in G\).

Assume to have a set of \(M\), RES-distributed, observation vectors \(x_{m} \in \ell_{m=1}^{M}\). Let us define \(\{\hat{x}_m\}_{m=1}^{M}\) as the set of \(M\)
vectors such that:
\[ \bar{x}_m = x_m - \hat{\mu}, \quad m = 1, \ldots, M \]
and \( \hat{\mu} \) is the sample mean estimator, i.e.,
\[ \hat{\mu} = N^{-1} \sum_{m=1}^{M} x_m. \]
The CSCM estimate can then be expressed as [42]:
\[ \hat{\Sigma}_{CSCM} = \frac{N}{\text{tr}((\hat{\Sigma}_{SCM})^{-1})} \Sigma_{SCM}, \]
\[ \Sigma_{SCM} = \frac{1}{M} \sum_{m=1}^{M} \bar{x}_m \bar{x}_m^T, \]
while C-Tyler and C-Hub estimates are the convergence points of the following iterative algorithm:
\[ S_T^{(k+1)} = \frac{1}{M} \sum_{m=1}^{M} \varphi(t(k)) \bar{x}_m \bar{x}_m^T, \]
\[ \varphi(t(k)) = NS_T^{(k+1)}/\text{tr}(S_T^{(k+1)}), \]
where \( t(k) = \bar{x}_m \Sigma_T^{(k+1)^{-1}} \bar{x}_m \) and the starting point is \( \Sigma_T^{(0)} = \mathbf{I} \).
The weight function \( \varphi(t) \) for Tyler’s estimator is defined as (see e.g., [14], [15, Ch. 4], and [43] and references therein):
\[ \varphi_{Tyler}(t) = N/t, \]
whereas the weight function for Huber’s estimator is given by ([14], [15, Ch. 4], [44]):
\[ \varphi_{Huber}(t) = \begin{cases} 1/b & t \leq \delta^2 \\ \delta^2/(tb) & t > \delta^2 \end{cases}, \]
and \( \delta = F_{\chi^2}^{-1}(u) \) where \( F_{\chi^2}(\cdot) \) indicates the distribution of a chi-squared random variable with \( N \) degrees of freedom and \( u \in (0,1) \) is a tuning parameter. Moreover, the parameter \( b \) is usually chosen as \( b = F_{\chi^2}^{-1}(\delta^2) + \delta^2(1 - F_{\chi^2}^{-1}(\delta^2))/N \) [14], [44].

To compare the three scatter matrix estimators with the CRSB, we define the following performance index:
\[ \varepsilon_{\alpha} = \| E\{ \text{vec}(\Sigma_B) - \text{vec}(\Sigma_0) \} \|_F, \]
where \( \alpha = \{ \text{CSCM, C-Tyler, C-Hub}(u) \} \) indicates the particular estimator under test and similarly \( \| A \|_F^2 = \text{tr}(A^TA) \) is the Frobenius norm of matrix \( A \). Similarly, for the sample mean, we have the following index:
\[ \varepsilon_{\mu_0} = \| E\{ (\hat{\mu} - \mu_0)(\hat{\mu} - \mu_0)^T \} \|_F. \]

For the sake of comparison, we show in the following figures also the constrained CRB (CCRB) for the estimation of \( \phi_0 = (\mu_0^T, \text{vec}(\Sigma_0)^T)^T \in \Omega \). The classical FIM can be obtained from the score vectors previously derived in Subsection IV-A as a block matrix of the form [39]:
\[ I(\phi_0) = E_0\{ s_{\phi_0} s_{\phi_0}^T \} = \begin{pmatrix} C_0(s_{\mu_0}) & 0 \\ 0^T & C_0(s_{\text{vec}(\Sigma_0)}) \end{pmatrix}. \]

Through direct calculation, it is easy to verify that:
\[ C_0(s_{\mu_0}) = C_0(s_{\mu_0}) = \frac{4E\{ Q\psi_0(Q)^2 \}}{N} \Sigma_0^{-1}, \]
\[ C_0(s_{\text{vec}(\Sigma_0)}) = \mathbf{D}_N^T \left( a_1 \text{vec}(\Sigma_0^{-1}) \text{vec}(\Sigma_0^{-1})^T + a_2 \Sigma_0^{-1} \otimes \Sigma_0^{-1} \right) \mathbf{D}_N \]
where:
\[ a_1 \triangleq \frac{1}{4} + \frac{E\{ Q\psi_0(Q) \}^2}{N} + \frac{a_2}{2}, \]
\[ a_2 \triangleq \frac{2E\{ Q^2\psi_0(Q)^2 \}}{N(N+2)}. \]

It is worth highlighting here that the difference between the classical FIM in (67) and the SFIM in (54) is due to the different expressions of the score vector \( s_{\text{vec}(\Sigma_0)} \) in (56) and of the one of the score vector \( s_{\text{vec}(\Sigma_0)} \) in (69).

Finally, the CCRB on the estimation of \( \phi_0 \in \Omega \) can be obtained using exactly the same procedure discussed for the CRSB in Subsection IV-D (see also [41]).

As performance bounds, the following indices are plotted:
\[ \varepsilon_{CSCM,\phi_0} = ||\text{CSCRB}(\phi_0)||_F, \]
\[ \varepsilon_{CSCM,\mu_0} = ||\text{CSCRB}(\phi_0, \mu_0)||_F, \]
\[ \varepsilon_{CSCM,\Sigma_0} = ||\text{CSCRB}(\phi_0, \mu_0)||_{\Sigma_0}, \]
where \( [,]_{\Sigma_0} \) and \( [,]_{\Sigma_0} \) indicate the top-left and the bottom-right submatrices of the CRB(\phi_0, \mu_0) of the CCRB(\phi_0, \mu_0).

We analyze two different cases:
1) The true RES distribution is a t-distribution.
2) The true RES distribution is a Generalized Gaussian (GG) distribution.

The simulation parameters that are common to the two cases are:
- \( [\Sigma_0]_{i,j} = \rho^{|i-j|}, \ i,j = 1, \ldots, N \). Moreover, \( \rho = 0.8 \) and \( N = 8 \).
- The data power is chosen to be \( \sigma^2 = E_0\{ Q \} / N = 4 \).
- The data mean value is chosen to be \( [\mu_0]_{i} = 1, \ i = 1, \ldots, N \).
- The number of the available independent and identically distributed (i.i.d.) data vectors is \( M = 3N = 24 \). Note that, since we assume to have \( M \) i.i.d. data vectors, the SFIM in (54) and the FIM in (67) have to be multiply by \( M \).
- The tuning parameter \( u \) of Huber’s estimator has been chosen as \( u = 0.9, 0.5, 0.1 \). Note that for \( u = 1 \) Huber’s estimator is equal to the SCM, while for \( u \to 0 \) Huber’s estimator tends to Tyler’s estimator [14].
- The number of independent Monte Carlo runs is \( 10^6 \).

A. Case 1: The t-Distribution
The density generator for the t-distribution is:
\[ g_t(t) \triangleq \frac{2^{N/2}}{\pi^{N/2} \Gamma(\lambda/2)} \left( \frac{\lambda}{\eta} \right)^{\lambda/2} \left( \frac{\lambda}{\eta} + t \right)^{\lambda/2 - N/2}. \]
and then \( \psi_0(t) \) in (36) is given by \( \psi_0(t) = -2^{-1}(\lambda + N)(\lambda + t)^{-1} \). Consequently, from (13), we have that:

\[
p_{Q}(q) = \frac{\Gamma\left(\frac{\lambda + N}{2}\right)}{\Gamma(\lambda/2)\Gamma(N/2)} \left(\frac{\lambda}{\eta}\right)^{\lambda/2 - 1} \left(\frac{\lambda}{\eta} + q\right)^{-\lambda N/2}.
\]

Using the integral in [45, pp. 315, n. 3.194 (3)], we have that:

\[
E\{Q\psi_0(Q)\} = -N/2
\]

\[
E\{Q\psi_0(Q)^2\} = \frac{\eta N(\lambda + N)}{4(\lambda + N + 2)}
\]

\[
E\{Q^2\psi_0(Q)^2\} = \frac{N(\lambda + N)(\lambda + N)}{4(\lambda + N + 2)}
\]

Then, the coefficients (70) and (71) for the \( t \)-distribution are:

\[
a_{1,t} \triangleq -\frac{1}{2(N + \lambda + 2)}, \quad a_{2,t} \triangleq \frac{\lambda + N}{2(N + \lambda + 2)}.
\]

By substituting the previous results in (56), (68), and (69), we obtain the matrices \( C_0(s_{\mu_0}), C_0(s_{\vec{v}e}(\Sigma_0)), \) and \( C_0(s_{\vec{v}e}(\Sigma_0)) \) for the \( t \)-distribution.

Fig. 1 shows the MSE index of the sample mean compared with the CSCRB as function of the shape parameter \( \lambda \). As we can see, when \( \lambda \to \infty \), the sample mean tends to be an efficient estimator. This is an expected result, since, when the shape parameter \( \lambda \) goes to infinity, the \( t \)-distribution becomes a Gaussian distribution, then the sample mean is the ML estimator for \( \mu_0 \). In Fig. 2 we can observe an interesting fact: the distance between the CSCRB and the CCRB increases as \( \lambda \to \infty \). This means that the lack of knowledge of the particular density generator, i.e., the lack of knowledge of the particular RES distribution of the data, has an higher impact when the tails of the true distribution become lighter. This behavior has been already observed in [20]. Regarding the \textit{constrained} scatter matrix estimators, the CSCM achieves the CSCRB as \( \lambda \to \infty \), i.e., as the data tends to be Gaussian distributed. Note that while it is well-known that the SCM is the ML estimator for the \textit{unconstrained} scatter matrix, the CSCM is \textit{not} the ML estimator for the constrained scatter matrix \( \Sigma_0 \). This is why, as \( \lambda \to \infty \), the CSCM does not achieve the CCRB. Regarding C-Tyler’s and C-Huber’s estimators, from Fig. 2 we can see that C-Huber’s estimator has better estimation performance for all the three values of \( u \) with respect to C-Tyler’s estimator. In particular, the estimation performance of C-Huber’s estimator improves as \( u \) tends to 1, i.e., when \( \lambda \) tends to collapse to the CSCM. Both C-Tyler’s and C-Huber’s estimators are far from being efficient with respect to the CSCRB. However, it must be mentioned here that efficiency is not the only property that an estimator should have. Robustness is also important in the choice of an estimation algorithm. We will investigate the trade-off between efficiency and robustness in future works.

Another important question that may arise is related to the behaviour of the Maximum Likelihood estimator of the scatter matrix. To answer this question, firstly we have to note that the density generator of the \( t \)-distribution in (75) depends on two additional parameters: the shape \( \lambda \) and the scale \( \eta \). If we assume to know perfectly both the functional form of \( g_0 \) and the scale and shape parameters, then the ML estimator of the scatter matrix \( \Sigma \) will outperform all \( M \)-estimators and will achieve the classical CRB for \( \Sigma \). This scenario is discussed in [39]. A more realistic situation is when only the functional form of \( g_0 \) is assumed to be a priori known while the scale and shape parameters have to be jointly estimated with the scatter matrix \( \Sigma \). A joint ML (JML) estimator for \( \Sigma \), \( \lambda \), and \( \eta \) generally does not exists, so we have to rely of sub-optimal strategies as the one discussed in [41]. Specifically, the recursive joint estimator of \( \Sigma \), \( \lambda \), and \( \eta \) proposed in [41] exploits the Method of Moments (MoM) to estimate the scale and shape parameters, while an estimation of the scatter matrix is obtained using an ML approach where the unknown parameters \( \lambda \) and \( \eta \) are replaced by their MoM estimates. Clearly, this JML algorithm will no longer achieve the classical CRB on the scatter matrix due to the lack of a priori knowledge about the scale and shape parameters. It would be interesting to investigate how the JML estimator behaves with respect to the CSCRB. As we can see from Fig. 1, the MSE index of the JML is larger than the CSCRB and this would suggest that not knowing the shape and scale parameters has the same impact of not knowing the whole functional form of the density generator. Of course, this aspect deserves further investigation and we leave it to future work.
The CSCM is an efficient estimator for \( \psi^{-1} s > = 3 + 2 N / \Sigma \) when the data is Gaussian distributed. In fact, as we move w.r.t. the \( s_2^2 \Gamma(2a) \), i.e., when the Gaussian distribution becomes lighter tails compared to the Gaussian distribution that can be obtained for \( s = 1 \).

The CSCRB when the data is Gaussian distributed, i.e., when \( s = 1 \) (see Fig. 4). However, it does not achieve the CCRB since, as discussed before, the CSCM is not the ML estimator for the constrained scatter matrix under Gaussian distributed data.

\[ g_0(t) \triangleq \frac{2^{N/2} s \Gamma(N/2)}{\pi^{N/2} (2b)^{N/2} s \Gamma(N/2s)} \exp \left( -\frac{t^2}{2b} \right) \]  

and then \( \psi_0(t) \) in (36) is given by \( \psi_0(t) = -s(2b)^{-1} t^{-1} \).

Consequently, from (13), we have that:

\[ p_q(q) = \frac{sq^{N/2-1}}{(2b)^{N/2} s \Gamma(N/2s)} \exp \left( -\frac{q^2}{2b} \right). \]

Using the integral in [45, pp. 370, n. 3.478 (1)], we have that:

\[ E\{Q\psi_0(Q)\} = -N/2. \]

\[ E\{Q^2\psi_0(Q)^2\} = \frac{s^2 \Gamma(2 + 4s - 2s)}{(2b)^{1/\Sigma} s \Gamma(N/2s)} \]

\[ E\{Q^2\psi_0(Q)^2\} = N(N + 2s)/4. \]

Then, the coefficients (70) and (71) for the GG distribution are:

\[ a_{1,GG} \triangleq \frac{s - 1}{2(N + 2)}, \quad a_{2,GG} \triangleq \frac{N + 2s}{2(N + 2)}. \]

As before, by substituting the previous results in (56), (68), and (69), we obtain the matrices \( C_0(s_{\mu_0}), \ C_0(s_{vecs}(\Sigma_0)) \) and \( C_0(s_{vecs}(\Sigma_1)) \) for the GG distribution.

The simulation results for the GG distributed data confirm all the previous discussions about \( t \)-distributed data:

- The sample mean estimator is an efficient estimator of \( \mu_0 \) when the data is Gaussian distributed. In fact, as we can see from Fig. 3 that the MSE index \( \varepsilon_{\mu_0} \) equates the CSCRB for \( s = 1 \), i.e., when the GG distribution becomes the Gaussian one.
- The distance between the CSCRB and the CCRB increases as the tails of the true data distribution become lighter. This behavior can be observed in Fig. 4. It is worth recalling here that for \( 0 < s < 1 \) the GG distribution has heavier tails and for \( s > 1 \) lighter tails compared to the Gaussian distribution that can be obtained for \( s = 1 \).
- The CSCM is an efficient estimator for \( \Sigma_0 \) w.r.t. the CSCRB when the data is Gaussian distributed, i.e., when \( s = 1 \).
background material for the derivation of the CSCRB provided in this paper. Let us introduce the underlying probability space \((X, \mathcal{F}, P_X)\), where \(X \subseteq \mathbb{R}^N\) is the sample space, \(\mathcal{F}\) is the Borel \(\sigma\)-algebra of events in \(X\) and \(P_X\) is the probability measure. Let \(x \in X\) be a random vector, then \(P_X(a) \triangleq P_X(x_1 \leq a_1, \ldots, x_N \leq a_N)\) is its cumulative distribution function (cdf). We assume that the cdf \(P_X\) admits a relevant probability density function (pdf) (with respect to the standard Lebesgue measure), denoted as \(p_X\), such that \(dP_X(a) = p_X(a) da\).

Consider now the vector space of the one-dimensional square-integrable function defined on \((X, \mathcal{F}, P_X)\):

\[
L_2(P_X) = \left\{ h : X \rightarrow \mathbb{R} \left| \int_X |h(x)|^2 dP_X(x) < \infty \right. \right\},
\]

with the inner product given by:

\[
\langle h, g \rangle_X = \int_X (h(x)g(x)dP_X(x), \forall h, g \in L_2(P_X).
\]

Let us define \(\mathcal{H}^q \subseteq L_2(P_X)\) as the subspace of all the one-dimensional, zero-mean functions on \((X, \mathcal{F}, P_X)\) such that:

\[
\mathcal{H}^q = \{ h \in L_2(P_X) | E_X[h] = 0 \}.
\]

It is immediate to verify that \(\mathcal{H}^q\), endowed with the inner product defined in (88), is an infinite-dimensional Hilbert space [27, Ch. 2].

The “\(q\)-replicating” Hilbert space \(\mathcal{H}^q\) of the zero-mean, \(q\)-dimensional functions on \((X, \mathcal{F}, P_X)\) is defined as the Cartesian product of \(q\) copies of the Hilbert space \(\mathcal{H}^1\), i.e., \(\mathcal{H}^q \triangleq \mathcal{H}^1 \times \mathcal{H}^1 \times \cdots \times \mathcal{H}^1\) [27, Ch. 3, Def. 6], such that:

\[
\mathcal{H}^q = \{ h = [h_1, \ldots, h_q]^T | h_i \in \mathcal{H}^1, i = 1, \ldots, q \}.
\]

Due to the Cartesian product-based construction, the inner product of \(\mathcal{H}^q\) is naturally induced by the one of \(\mathcal{H}^1\), i.e.,:

\[
\langle h_1, h_2 \rangle_X \triangleq E_X[h_1^T h_2] = \sum_{i=1}^q E_X[h_{1,i}h_{2,i}],
\]

and consequently the norm is:

\[
||h||_X = \sqrt{\langle h, h \rangle_X} = \sqrt{\sum_{i=1}^q E_X[h_i^2]}.
\]

Let us now investigate the geometrical structure of \(\mathcal{H}^q\), with a particular focus on the orthogonal projection of a generic element \(h \in \mathcal{H}^q\) into a closed subspace \(\mathcal{V}\) of \(\mathcal{H}^q\). The following theorem is a fundamental result in Hilbert spaces theory, and can be established in a very general setting (see e.g., [46, Th. 3.9.3]). Here, we will adapt it to the particular Hilbert space \(\mathcal{H}^q\).

**Theorem VI.1 (The Projection Theorem):** Let \(\mathcal{V}\) be a closed subspace of the Hilbert space \(\mathcal{H}^q\) and let \(h\) and \(g\) be two \(q\)-dimensional, zero-mean functions on \((X, \mathcal{F}, P_X)\), such that \(h \in \mathcal{H}^q\) and \(g \in \mathcal{V}\). Then, the following conditions are equivalent:

1) \(||h - g||_X = \inf_{h' \in \mathcal{V}} ||h - h'||_X\),
2) \(h\) can be uniquely written as

\[
h = g + (h - g),
\]

where \(g \triangleq \Pi(h|\mathcal{V}) \in \mathcal{V}\) and \(h - \Pi(h|\mathcal{V}) \in \mathcal{V}^\perp\), where \(\mathcal{V}^\perp\) indicates the orthogonal complement of \(\mathcal{V}\) in \(\mathcal{H}^q\),

3) the element \(g \triangleq \Pi(h||\mathcal{V}) \in \mathcal{V}\) is then uniquely determined by the orthogonality constraint

\[
(h - \Pi(h|\mathcal{V}), 1)_X = 0, \quad \forall \mathcal{V}.
\]

The operator \(\Pi(\cdot|\mathcal{V})\) defined in Theorem VI.1 is called the orthogonal projection operator onto the closed subspace \(\mathcal{V}\). The unique element \(\Pi(h|\mathcal{V})\) is then called the orthogonal projection of \(h\) in \(\mathcal{H}^q\) onto \(\mathcal{V}\). Furthermore, as a consequence of the Condition 2) in Theorem VI.1, the Hilbert space \(\mathcal{H}^q\) can be written as the Cartesian product of the subspace \(\mathcal{V}\) and of its orthogonal complement \(\mathcal{V}^\perp\), i.e., \(\mathcal{H}^q = \mathcal{V} \times \mathcal{V}^\perp\).

**APPENDIX B**

**PROJECTION OPERATOR AND CONDITIONAL EXPECTATION**

In Appendix A we defined \(\mathcal{H}^q\) as the Hilbert space of the \(q\)-dimensional, zero-mean functions on the probability space \((X, \mathcal{F}, P_X)\). Let \(\mathcal{G}(\mathcal{V}) \subseteq \mathcal{F}\) be the sub-sigma algebra generated by the random variable \(V\). It can be shown (see e.g., [5, Appendix 3] and [47, Ch. 23]) that the set of all the \(q\)-dimensional, zero-mean functions on the probability space \((X, \mathcal{G}(\mathcal{V}), P_X|\mathcal{V})\) is a closed linear subspace, say \(\mathcal{V}\), of the Hilbert space \(\mathcal{H}^q\).

This fact can be exploited to establish a link between the projection operator \(\Pi(\cdot|\mathcal{V})\) and the conditional expectation \(E_X|\mathcal{V}(\cdot|\mathcal{V})\). Firstly, let us define \(E_X|\mathcal{V}(\cdot|\mathcal{V})\) as in [5, Appendix 3] and [47, Ch. 23].

**Definition VII.1:** Let \(h \in \mathcal{H}^q\) and \(I \in \mathcal{V} \subseteq \mathcal{H}^q\) be two \(q\)-dimensional, zero-mean functions on the probability spaces \((X, \mathcal{F}, P_X)\) and \((X, \mathcal{G}(\mathcal{V}), P_X|\mathcal{V})\) with \(\mathcal{G}(\mathcal{V}) \subseteq \mathcal{F}\), respectively. Then the conditional expectation \(E_X|\mathcal{V}(h|\mathcal{V})\) is the unique element in \(\mathcal{V}\), such that:

\[
(h - E_X|\mathcal{V}(h|\mathcal{V}), 1)_X \triangleq E_X(h - E_X|\mathcal{V}(h|\mathcal{V})) = 0,
\]

for every \(I \in \mathcal{V}\).

For a more general and formal definition we refer the reader to [5, Appendix 3] and [47, Ch. 23]. The condition (95) is equivalent to (94) in Theorem VI.1 that defines the projection operator, and consequently, we have that:

\[
\Pi(\cdot|\mathcal{V}) = E_X|\mathcal{V}(\cdot|\mathcal{V}).
\]

The usefulness of this relation is the fact that, for some semiparametric models, the tangent space presents an invariance structure with respect to a group of transformations and it admits a characterization through a certain sub-sigma algebra generated by the relevant maximal invariant statistic [30]. An example of a semiparametric model that owns this property is the semiparametric group model of RES distributions discussed in Section III-B.

**REFERENCES**

[1] 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 Process. Mag., vol. 34, no. 6, pp. 142–157, Nov. 2017.

[2] S. Fortunati, M. S. Greco, and F. Gini, “Asymptotic robustness of Kelly’s GLRT and Adaptive Matched Filter detector under model misspecification,” in Proc. ISI World Statist. Congr., 2017. [Online]. Available: https://arxiv.org/abs/1709.08667
[3] P. J. Bickel and J. Kwon, “Inference for semiparametric models: Some questions and an answer,” Statist. Sinica, vol. 11, pp. 863–960, 2001.

[4] J. A. Wellner, “Semiparametric models: Progress and problems,” Bull. Int. Statist. Inst., vol. 51, pp. 2–24, 1985.

[5] P. Bickel, C. Klaassen, Y. Ritov, and J. Wellner, Efficient and Adaptive Estimation for Semiparametric Models. Baltimore, MD, USA: Johns Hopkins Univ. Press, 1993.

[6] S.-I. Amari and J. F. Cardoso, “Blind source separation-semiparametric statistical approach,” IEEE Trans. Signal Process., vol. 49, no. 11, pp. 2692–2700, Nov. 1997.

[7] H. Hammes, E. Woisetsky, and A. M. Zoubir, “Transformation-based robust semiparametric estimation,” IEEE Signal Process. Lett., vol. 15, pp. 845–848, 2008.

[8] S. Cambanis, S. Huang, and G. Simons, “On the theory of elliptically contoured distributions,” J. Multivariate Anal., vol. 11, no. 3, pp. 368–385, 1981.

[9] C. D. Richmond, “Adaptive array signal processing and performance analysis in non-Gaussian environments,” Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, MA, USA, 1996. [Online]. Available: https://dspace.mit.edu/handle/1721.1/11005

[10] M. A. Chmielewski, “Elliptically symmetric distributions: A review and bibliography,” Int. Statist. Rev., vol. 49, no. 1, pp. 67–74, 1981.

[11] K. Yao, “A representation theorem and its applications to spherically-invariant random processes,” IEEE Trans. Inf. Theory, vol. 19, no. 5, pp. 600–608, Sep. 1973.

[12] S. Zozor and C. Vignat, “Some results on the denoising problem in the elliptically distributed context,” IEEE Trans. Signal Process., vol. 58, no. 1, pp. 134–150, Jan. 2010.

[13] M. Greco, S. Fortunati, and F. Gini, “The misspecified Cramér-Rao bound,” IEEE Trans. Signal Process., vol. 60, no. 11, pp. 5597–5625, Nov. 2012.

[14] A. M. Zoubir, V. Koivunen, E. Ollila, and M. Muma, Robust Statistics for Signal Processing. Cambridge, UK: Cambridge Univ. Press, 2018.

[15] S. Fortunati, F. Gini, M. S. Greco, A. M. Zoubir, and M. Ragasawamy, “A fresh look at the semiparametric Cramér-Rao bound,” Proc. 26th Eur. Signal Process. Conf., 2018. [Online]. Available: http://arxiv.org/abs/1803.00267

[16] P. J. Bickel, “On adaptive estimation,” Annu. Statist., vol. 10, no. 3, pp. 647–671, 1982.

[17] M. Hallin and D. Paindaveine, “Semiparametrically efficient rank-based inference for shape I. Optimal rank-based tests for sphericity,” Ann. Statist., vol. 34, no. 6, pp. 2707–2756, 2006.

[18] M. Hallin, H. Oja, and D. Paindaveine, “Semiparametrically efficient rank-based inference for shape II. Optimal re-estimation of shape,” Ann. Statist., vol. 34, no. 6, pp. 2757–2789, 2006.

[19] M. Hallin and D. Paindaveine, “Parametric and semiparametric inference for shape: The role of the scale functional,” Statist. Decis., vol. 24, no. 3, pp. 327–350, 2009.

[20] D. Paindaveine, “A canonical definition of shape,” Statist. Probab. Lett., vol. 78, no. 14, pp. 2240–2247, 2008.

[21] L. LeCam and G. L. Yang, Asymptotics in Statistics: Some Basic Concepts (Springer Series in Statistics), 2nd ed., New York, NY, USA: Springer, 1990.

[22] J. A. Wellner, “Semiparametric models: Progress and problems,” Bull. Int. Statist. Inst., vol. 51, pp. 2–24, 1985.

[23] W. K. Newey, “The asymptotic variance of semiparametric estimators,” Econometrica, vol. 62, no. 6, pp. 1349–1382, 1994.

[24] C. A. J. Klaassen, “Consistent estimation of the influence function of locally asymptotically linear estimators,” Ann. Statist., vol. 15, no. 4, pp. 1548–1562, Dec. 1987.

[25] A. Tsitsi, Semiparametric Theory and Missing Data (Springer Series in Statistics). New York, NY, USA: Springer, 2000.

[26] H. Rieder, Robust Asymptotic Statistics (Springer Series in Statistics). New York, NY, USA: Springer, 1994.

[27] K.-T. Fang, S. Kotz, and K. W. Ng, Symmetric Multivariate and Related Distributions (Monographs on Statistics and Applied Probability), New York, NY, USA: Springer, 1990.

[28] E. L. Lehmann and J. P. Romano, Testing Statistical Hypotheses (Springer Texts in Statistics). New York, NY, USA: Springer, 2004.

[29] J. R. Magnus and H. Neudecker, “The commutation matrix: Some properties and applications,” Ann. Statist., vol. 7, no. 2, pp. 381–394, Mar. 1979.
Fulvio Gini (’F06) received the Doctor Engineer (cum laude) and Research Doctor degrees in electronic engineering from the University of Pisa, Pisa, Italy, in 1990 and 1995, respectively. In 1993, he joined the Department of Electrical Engineering, University of Virginia, Charlottesville, VA, USA. He has been an Associate Editor for the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS since January 2007 and for Signal Processing since December 2006. He was an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2000 to 2006 and has been a Senior Associate Editor for the same Transactions since February 2016. He was a member of the Editorial Board of the EURASIP Journal on Advances in Signal Processing. He was a co-founder and first co-Editor-in-Chief of the Hindawi International Journal on Navigation and Observation from 2007 to 2011. He was the Area Editor for the Special issues of the IEEE SIGNAL PROCESSING MAGAZINE from 2012 to 2014. He was a co-recipient of the 2001 and 2012 IEEE Aerospace and Electronic Systems (AES) Society’s Barry Carlton Award for the Best Paper published in the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS. He was a recipient of the 2003 IEEE Achievement Award for outstanding contribution in signal processing and of the 2003 IEEE AES Society Nathanson Award to the Young Engineer of the Year. He is a member of the Radar System Panel (from 2008 to present) and of the Board of Governors (from 2017 to 2019) of the IEEE Aerospace and Electronic Systems Society. He is a member of the IEEE Signal Processing Society Awards Board (2016–2018). He has been a member of the Signal Processing Theory and Methods Technical Committee (TC) of the IEEE Signal Processing Society and of the Sensor Array and Multichannel TC for many years. He is a member of the Board of Directors of the European Association for Signal Processing (EURASIP), where he was the Award Chair (2006–2012) and the EURASIP President (2013–2016). He is the General Co-Chair of the 2020 IEEE Radar conference to be held in Florence, Italy, in September 2020. He was the Technical Co-Chair of the 2006 EURASIP Signal and Image Processing Conference, Florence, of the 2008 Radar Conference, Rome, Italy, and of the 2015 IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, Cancún, Mexico. He was the General Co-Chair of the 2nd Workshop on Cognitive Information Processing, 2010, of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing, Florence, and of the second, third, and fourth editions of the Workshop on Compressive Sensing in Radar. He was the Section Editor for the Radar Signal Processing section, vol. 3 of the Academic Press Library in Signal Processing (S. Theodoridis and R. Chellappa, Eds., New York, NY, USA: Elsevier, 2013). He is a co-editor and author of the books entitled Knowledge Based Radar Detection, Tracking and Classification (2008) and Waveform Diversity and Design (2012). He authored or coauthored more than 130 journal and 160 conference papers. He was a Guest Co-Editor of two special sections of the Journal of the IEEE Signal Processing Society on Special Topics in Signal Processing, one on “Adaptive Waveform Design for Agile Sensing and Communication” (2007) and the other on “Advanced Signal Processing for Time-Frequency Modulated Arrays” (2017), a Guest Editor of the special section of the IEEE SIGNAL PROCESSING MAGAZINE on “Knowledge Based Systems for Adaptive Radar Detection, Tracking and Classification” (2006), and a Guest Co-Editor of the two special issues of the EURASIP Signal Processing journal on “New trends and findings in antenna array processing for radar” (2004) and on “Advances in Sensor Array Processing (in memory of A. Gershman)” (2013). December 2016. In May–June 2015 and in January–February 2018, she visited as an invited Professor the Université Paris-Sud, CentraleSupélec, Paris, France. She co-authored many book chapters and more than 190 journal and conference papers. Her general research interests include statistical signal processing, estimation, and detection theory, in particular clutter models, coherent and incoherent detection in non-Gaussian clutter, constant false alarm rate techniques, radar waveform diversity, bistatic/multistatic active and passive radars, and cognitive radars. She was a co-recipient of the 2001 and 2012 IEEE Aerospace and Electronic Systems Society’s Barry Carlton Awards for the Best Paper and a recipient of the 2008 Fred Nathanson Young Engineer of the Year award for contributions to signal processing estimation, and detection theory. She has been a General Chair, Technical Program Chair, and Organizing Committee Member of many international conferences over the past ten years. She is a Guest Editor of the upcoming special issue on Machine Learning for Cognition in Radio Communications and Radar of the IEEE JOURNAL ON SPECIAL TOPICS OF SIGNAL PROCESSING. She was a Lead Guest Editor of the special issue on “Advanced Signal Processing for Radar Applications” of the IEEE JOURNAL ON SPECIAL TOPICS OF SIGNAL PROCESSING in December 2015, a Guest Co-Editor of the special issue of the Journal of the IEEE Signal Processing Society on Special Topics in Signal Processing on “Adaptive Waveform Design for Agile Sensing and Communication,” published in June 2007, and a Lead Guest Editor of the special issue of the International Journal of Navigation and Observation on Modelling and Processing of Radar Signals for Earth Observation published in August 2008. She is an Associate Editor for IET Proceedings—Sonar, Radar and Navigation, a member of the Editorial Board of the Springer’s Journal of Advances in Signal Processing, a Senior Editorial Board Member of the IEEE JOURNAL ON SELECTED TOPICS OF SIGNAL PROCESSING, and a Senior Area Chair for the IEEE TRANSACTIONS ON SIGNAL PROCESSING. She was a member of the IEEE Aerospace and Electronic Systems Society (AESS) and IEEE Signal Processing Society Board of Governors (2015–2017) and the Past Chair of the IEEE AESS Radar Panel. She was an IEEE Signal Processing Society Distinguished Lecturer for the years 2014–2015. She is an IEEE AESS Distinguished Lecturer for the years 2015–2018 and a member of the IEEE Fellow Committee and AESS VP Publications.

Abdelrahman M. Zoubir (’F08) received the Dr.-Ing. degree from Ruhr-Universität Bochum, Bochum, Germany, in 1992. From 1992 to 1998, he was with the Queensland University of Technology, Australia, where he was an Associate Professor. In 1999, he joined the Curtin University of Technology, Australia, as a Professor of telecommunications and was the Interim Head of the School of Electrical and Computer Engineering from 2001 until 2003. In 2003, he joined Technische Universität Darmstadt, Darmstadt, Germany, as a Professor of signal processing and the Head of the Signal Processing Group. He authored or coauthored more than 400 journal and conference papers on these areas. His research interests include statistical methods for signal processing with an emphasis on bootstrap techniques, robust detection and estimation, and array processing applied to telecommunications, radar, sonar, car engine monitoring and safety, and biomedicine. He was the General or Technical Chair of numerous international workshops and conferences, notably the Technical Chair of the IEEE International Conference on Acoustics, Speech and Signal Processing, Florence, Italy, 2014. He was also on several editorial boards and was a Guest Editor on numerous occasions, notably the Editor-In-Chief of the IEEE SIGNAL PROCESSING MAGAZINE from 2012 until 2014. He was the Chair (2010–2011), Vice-Chair (2008–2009), and member (2002–2008) of the IEEE Signal Processing Society (SPS) Technical Committee Signal Processing Theory and Methods, and a member (2007–2012) of the IEEE SPS Technical Committee Sensor Array and Multi-Channel Signal Processing. He was a Guest Editor of four special issues of the IEEE JOURNAL ON SPECIAL TOPICS OF SIGNAL PROCESSING. He was a Member-at-Large (2015–2017). He was an elected member of the Board of Directors of the European Association of Signal Processing (EURASIP) from 2008 to 2016. He is the President of the EURASIP for 2017–2018. He was an IEEE Distinguished Lecturer (class 2010–2011).

Maria Sabrina Greco (’F11) received the Graduate degree in electronic engineering in 1993 and the Ph.D. degree in telecommunication engineering in 1998, both from the University of Pisa, Pisa, Italy. From December 1997 to May 1998, she was a visiting researcher Scholar with the Georgia Tech Research Institute, Atlanta, GA, where she conducted research activity in the field of radar detection in non-Gaussian background. In 1993, she joined the Department of Information Engineering, University of Pisa, where she has been a Full Professor since...
Muralidhar Rangaswamy (S’89–M’93–SM’98–F’06) received the B.E. degree in electronics engineering from Bangalore University, Bangalore, India, in 1985, and the M.S. and Ph.D. degrees in electrical engineering from Syracuse University, Syracuse, NY, USA, in 1992. He is currently a Senior Advisor for Radar Research at the RF Exploitation Branch within the Sensors Directorate of the Air Force Research Laboratory (AFRL), Wright-Patterson AFB, OH, USA. Prior to this, he has held industrial and academic appointments. His research interests include radar signal processing, spectrum estimation, modeling non-Gaussian interference phenomena, and statistical communication theory. A common theme of his research in these areas is that of big data. A large class of radar signal processing problems incurs the curse of dimensionality, i.e., the training data requirement and computational cost grow explosively with increasing dimensionality. He has developed innovative techniques to address these challenges by using physics-based approaches and exploiting underlying structure in the data gained from a priori information. He has coauthored more than 200 refereed journal and conference record papers in the areas of his research interests. Additionally, he has contributed to eight books and coinvented three U.S. patents. He served as the Technical Editor (Associate Editor-in-Chief) for Radar Systems in the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS (IEEE TAES) from 2007 to 2015 and as Associate Editor for Radar Systems within the IEEE TAES from 2004 to 2007. He was the co-Editor-in-Chief for the Digital Signal Processing journal between 2005 and 2011. He was a member of the Senior Editorial Board of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING from January 2012 to December 2014. He was a two-term elected member of the Sensor Array and Multichannel Processing Technical Committee of the IEEE Signal Processing Society between January 2005 and December 2010 and a member of the Radar Systems Panel (RSP) in the IEEE Aerospace and Electronic Systems Society (AESS). He was the General Chairman for the Fourth IEEE Workshop on Sensor Array and Multichannel Processing, Waltham, MA, USA, July 2006. He has served on the Technical Committee of the IEEE Radar Conference series in a myriad of roles (Track Chair, Session Chair, Special Session Organizer and Chair, Paper Selection Committee Member, and Tutorial Lecturer). He served as the Publicity Chair for the First IEEE International Conference on Waveform Diversity and Design, Edinburgh, U.K., November 2004. He presently serves on the Conference Sub-Committee of the RSP. He was the Technical Program Chairman for the 2014 IEEE Radar Conference. He was a recipient of the IEEE Warren White Radar Award in 2013, the 2013 Affiliate Societies Council Dayton Outstanding Scientist and Engineer Award, the 2007 IEEE Region 1 Award, the 2006 IEEE Boston Section Distinguished Member Award, and the 2005 IEEE AESS Fred Nathanson Memorial Outstanding Young Radar Engineer Award. He was also a recipient of the 2012 and 2005 Charles Ryan Basic Research Award from the Sensors Directorate of AFRL, in addition to more than 40 scientific achievement awards.