PERFORMANCE ENHANCEMENT OF THE MEASURE-TRANSFORMED MUSIC ALGORITHM VIA MSE BASED OPTIMIZATION

Nir Halay and Koby Todros

Ben-Gurion University of the Negev

ABSTRACT

The measure-transformed (MT) MUltiple SIgnal Classification (MUSIC) algorithm is a robust MUSIC generalization that operates by applying a transform to the probability measure (distribution) of the data. In this paper, we first provide an asymptotic mean-squared-error (MSE) performance analysis of the MT-MUSIC algorithm. Under some mild assumptions, we show that the MT-MUSIC estimator is asymptotically normal and unbiased, and obtain an analytic expression for the asymptotic MSE matrix. We then proceed to develop a strongly consistent estimator for the asymptotic MSE matrix that is constructed from the same data samples being used for implementation of the MT-MUSIC. This paves the way for development of a data-driven procedure for optimal selection of the measure transformation parameters that minimizes an empirical estimate of the asymptotic average root MSE (RMSE). Simulation examples illustrate the performance advantage of the proposed MSE based optimization of the MT-MUSIC.

***Index Terms*—** Array processing, DOA estimation, probability measure transform, robust statistics, signal subspace estimation.

1. INTRODUCTION

The multiple signal classification (MUSIC) algorithm [1], [2] is a popular technique for estimating directions-of-arrival (DOAs) of noisy signals received by an array of sensors. This algorithm operates by finding DOAs with corresponding steering vectors that have minimal projections onto the empirical noise-subspace. The spanning vectors of the empirical noise subspace are obtained from the eigenvectors corresponding to the minimal eigenvalues of the sample covariance matrix (SCM).

The SCM is the maximum-likelihood estimator (MLE) of the covariance under the assumption that the data obeys a normal distribution [3]. This estimator is highly sensitive to large deviations from normality, that can occur, e.g., in the presence of heavy-tailed noise that generates outliers, resulting in poor DOAs estimates. To handle this problem, several robust MUSIC extensions have been proposed that replace the SCM with robust scatter matrix estimators. In [4], two robust MUSIC extensions were developed that replace the SCM with robust sign and rank covariance estimates. In [5], it was proposed to replace the non-robust SCM with robust M-estimators of scatter [6], such as the MLE (under elliptical observations), Huber's [7], and Tyler's [8]. In this context, we note that a consistent robust MUSIC extension based on M-estimators of scatter was developed in [9], [10] for the high-dimensional sample-starved scenario, which is not considered in this paper. A more elaborate discussion about the methods proposed in [4], [5], and other important robust MUSIC extensions [11]-[15], including their advantages and disadvantages, appears in [16, Sec. I].

Recently, we developed in [16] a robust MUSIC extension, called measure-transformed MUSIC (MT-MUSIC) that operates by applying a transform to the probability measure of the data. In a more specified manner, in MT-MUSIC the SCM is replaced by an empirical measure-transformed (MT) covariance. The considered measure transform, also applied in [17]-[22], is structured by a non-negative function, called MT-function, that weights the data samples. In [16], the MT-MUSIC was implemented with a Gaussian-shaped spherical MT-function parameterized by a width parameter. Under the Gaussian MT-function, we have shown that the empirical MT-covariance is B-robust [23], and, unlike other robust scatter matrix estimators, such as the empirical sign-covariance and Tyler's M-estimator of scatter, its influence function [23] decays to zero as the outlier norm approaches infinity, resulting in enhanced resilience against large-norm outliers. Additionally, under the assumption of spherical compound Gaussian (CG) noise [5], we proved that the noise subspace can be determined from the eigen-decomposition of the Gaussian MT-covariance. In [16], selection of the width parameter of the Gaussian MT-function was carried out via suboptimal procedure that controls the transform-domain Fisher information loss under nominal Gaussian distribution of the data.

Main contributions: This paper provides an important extension of the work presented in [16]. First, we analyze the asymptotic mean-squared-error (MSE) performance the MT-MUSIC algorithm. We emphasize that the analysis is not restricted to the Gaussian-shaped MT-function considered in [16]. Furthermore, we do not assume to a specific probability distribution of the data (e.g., Gaussian or elliptical). Under some mild regularity conditions, we show that the MT-MUSIC estimator is asymptotically normal and unbiased, and obtain an analytic expression for the asymptotic MSE matrix.

We go on to develop a strongly consistent estimator for the asymptotic MSE matrix. The MSE estimator is constructed via the same data samples used for implementation of the MT-MUSIC algorithm. This paves the way for development of a data-driven procedure for optimal selection of the MT-function within some parametric class. Unlike the suboptimal procedure considered in [16], the proposed selection procedure minimizes an empirical estimate of the asymptotic average root MSE (RMSE).

The proposed MSE based optimization of the MT-MUSIC is illustrated in simulation examples involving Gaussian and non-Gaussian heavy-tailed noise. Throughout these examples, the MT-MUSIC is implemented with a non-Gaussian outlier-suppressing MT-function whose parameter is selected via the optimal MSE based data-driven procedure discussed above. We show that the optimized non-Gaussian MT-MUSIC outperforms the suboptimal Gaussian MT-MUSIC [16] and other robust MUSIC extensions.

Paper organization: Section 2 provides a brief review of the MT-MUSIC algorithm [16]. In Section 3, we analyze the asymp-

This work was supported in part by United States-Israel Binational Science Foundation under Grant 2014334.

totic MSE performance of the MT-MUSIC and develop a data-driven procedure for optimal selection of the MT-function. In Section 4, the MSE based optimization of the MT-MUSIC is illustrated in simulation studies. Finally, section 5 provides concluding remarks. Proofs for the theorem and proposition stated in the paper will be provided in the full journal version.

2. MEASURE-TRANSFORMED MUSIC: REVIEW

We begin by introducing the considered sensor array model. Then, the relevant principles of the probability measure transform [16]-[22] are presented. Finally, the MT-MUSIC algorithm [16] is reviewed.

2.1. Data model

Consider an array of p sensors that receives signals from $q < p$ narrow-band far-field incoherent point sources with distinct azimuthal DOAs $\{\theta_1, ..., \theta_q\} \subset \Theta$, where $\Theta \subseteq [-\pi, \pi)$ denotes the parameter space. Under this setting, the array output satisfies the following observation model [2]:

$$\mathbf{X}_n = \mathbf{A}\mathbf{S}_n + \mathbf{W}_n, \quad n = 1, \ldots, N, \tag{1}$$

where $\{\mathbf{X}_n \in \mathbb{C}^p\}$ is an observation process, $\{\mathbf{S}_n \in \mathbb{C}^q\}$ is a first-order stationary latent signal process with zero-mean and nonsingular covariance, and $\{\mathbf{W}_n \in \mathbb{C}^p\}$ denotes a first-order stationary zero-mean spatially white noise processes that is independent of $\{\mathbf{S}_n\}$. The matrix $\mathbf{A} \triangleq [\mathbf{a}(\theta_1), ..., \mathbf{a}(\theta_q)] \in \mathbb{C}^{p\times q}$, is the array steering matrix, where $\mathbf{a}(\vartheta) \in \mathbb{C}^p$ is a steering vector toward direction $\vartheta \in \Theta$. The array is assumed to be unambiguous, i.e., any collection of p steering vectors corresponding to distinct DOAs forms a linearly independent set. Under this assumption the steering matrix $\mathbf{A}$ has a full column rank, and therefore, identification of its column vectors amounts to identification of the DOAs.

2.2. Probability measure transform

We define the measure space $(\mathcal{X}, \mathcal{S}_{\mathcal{X}}, P_{\mathbf{X}})$, where $\mathcal{X} \subseteq \mathbb{C}^p$ is the observation space of a random vector $\mathbf{X}$, $\mathcal{S}_{\mathcal{X}}$ is a σ-algebra over $\mathcal{X}$ and $P_{\mathbf{X}}$ is a probability measure on $\mathcal{S}_{\mathcal{X}}$.

Definition 1. *Given a non-negative function* $u : \mathbb{C}^p \to \mathbb{R}_+$ *satisfying* $0 < \mathrm{E}[u(\mathbf{X}); P_{\mathbf{X}}] < \infty$*, where* $\mathrm{E}[u(\mathbf{X}); P_{\mathbf{X}}] \triangleq \int_{\mathcal{X}} u(\mathbf{x})\, dP_{\mathbf{X}}(\mathbf{x})$*, a transform on* $P_{\mathbf{X}}$ *is defined as:*

$$Q^{(u)}_{\mathbf{X}}(A) \triangleq \mathbb{T}_u[P_{\mathbf{X}}](A) = \int_A \varphi_u(\mathbf{x})\, dP_{\mathbf{X}}(\mathbf{x}), \tag{2}$$

where $A \in \mathcal{S}_{\mathcal{X}}$ *and* $\varphi_u(\mathbf{x}) \triangleq u(\mathbf{x})/\mathrm{E}[u(\mathbf{X}); P_{\mathbf{X}}]$*. The function* $u(\cdot)$ *is called the MT-function.*

By [16, Prop. 1], $Q^{(u)}_{\mathbf{X}}$ is a probability measure on $\mathcal{S}_{\mathcal{X}}$ that is absolutely continuous w.r.t. $P_{\mathbf{X}}$, with Radon-Nikodym derivative [24] $dQ^{(u)}_{\mathbf{X}}(\mathbf{x})/dP_{\mathbf{X}}(\mathbf{x}) = \varphi_u(\mathbf{x})$. Thus, the covariance matrix of $\mathbf{X}$ under $Q^{(u)}_{\mathbf{X}}$ (the MT-covariance) is given by:

$$\mathbf{\Sigma}^{(u)}_{\mathbf{X}} \triangleq \mathrm{E}\left[\mathbf{X}\mathbf{X}^H \varphi_u(\mathbf{X}); P_{\mathbf{X}}\right] - \boldsymbol{\mu}^{(u)}_{\mathbf{X}} \boldsymbol{\mu}^{(u)H}_{\mathbf{X}}, \tag{3}$$

where $\boldsymbol{\mu}^{(u)}_{\mathbf{X}} \triangleq \mathrm{E}[\mathbf{X}\varphi_u(\mathbf{X}); P_{\mathbf{X}}]$ denotes the MT-mean. Equation (3) implies that the MT-covariance is a weighted covariance of $\mathbf{X}$ under $P_{\mathbf{X}}$, with the weighting function $\varphi_u(\cdot)$ defined below (2). Notice that when the MT-function $u(\cdot)$ is non-zero and constant valued, the standard covariance is obtained.

Given a sequence of samples $\mathbf{X}_1, \ldots, \mathbf{X}_N$ from $P_{\mathbf{X}}$, the empirical MT-covariance is defined as:

$$\hat{\mathbf{\Sigma}}^{(u)}_{\mathbf{X}} \triangleq \sum_{n=1}^{N} \mathbf{X}_n \mathbf{X}_n^H \hat{\varphi}_u(\mathbf{X}_n) - \hat{\boldsymbol{\mu}}^{(u)}_{\mathbf{X}} \hat{\boldsymbol{\mu}}^{(u)H}_{\mathbf{X}}, \tag{4}$$

where $\hat{\varphi}_u(\mathbf{X}_n) \triangleq u(\mathbf{X}_n)/\sum_{n=1}^{N} u(\mathbf{X}_n)$, and the empirical MT-mean $\hat{\boldsymbol{\mu}}^{(u)}_{\mathbf{X}} \triangleq \sum_{n=1}^{N} \mathbf{X}_n \hat{\varphi}_u(\mathbf{X}_n)$. According to [16, Prop. 2], if the samples are i.i.d. and $\mathrm{E}[\|\mathbf{X}\|^2 u(\mathbf{X}); P_{\mathbf{X}}] < \infty$ then $\hat{\mathbf{\Sigma}}^{(u)}_{\mathbf{X}}$ is a strongly consistent estimator of $\mathbf{\Sigma}^{(u)}_{\mathbf{X}}$.

Robustness of the empirical MT-covariance (4) to outliers was studied in [16] using its influence function [23], which describes the effect on the estimator introduced by an infinitesimal contamination at some point $\mathbf{y} \in \mathbb{C}^p$. An estimator is said to be B-robust if its influence function is bounded [23]. In [16, Prop. 3], it was shown that if there exists a finite positive constant M, such that

$$u(\mathbf{y}) \le M \ \text{ and } \ u(\mathbf{y})\|\mathbf{y}\|^2 \le M, \ \ \forall \mathbf{y} \in \mathbb{C}^p \tag{5}$$

then the influence function of $\hat{\mathbf{\Sigma}}^{(u)}_{\mathbf{X}}$ is bounded.

2.3. The MT-MUSIC algorithm

The MT-MUSIC [16] is comprised of two steps. First, the MT-function $u(\cdot)$ is chosen such that the following conditions are satisfied:
A-1) The resulting empirical MT-covariance $\hat{\mathbf{\Sigma}}^{(u)}_{\mathbf{X}}$ is B-robust.
A-2) Let $\lambda^{(u)}_1 \ge ... \ge \lambda^{(u)}_p$ denote the eigenvalues of the MT-covariance $\mathbf{\Sigma}^{(u)}_{\mathbf{X}}$. The $p-q$ smallest eigenvalues of $\mathbf{\Sigma}^{(u)}_{\mathbf{X}}$ satisfy

$$\lambda^{(u)}_q > \lambda^{(u)}_{q+1} = ... = \lambda^{(u)}_p \triangleq \gamma^{(u)} \tag{6}$$

and their corresponding eigenvectors span the null-space of $\mathbf{A}^H$, also called the noise-subspace.
Second, the DOAs are estimated by finding the q highest maxima of the measure-transformed pseudo-spectrum defined as:

$$\hat{P}^{(u)}(\vartheta) \triangleq \|\hat{\mathbf{G}}^{(u)H} \mathbf{a}(\vartheta)\|^{-2}, \tag{7}$$

where $\hat{\mathbf{G}}^{(u)} \in \mathbb{C}^{p\times p-q}$ denotes the matrix of $p-q$ eigenvectors of $\hat{\mathbf{\Sigma}}^{(u)}_{\mathbf{X}}$ corresponding to the smallest eigenvalues.

3. MSE BASED OPTIMIZATION OF THE MT-MUSIC

In this section we analyze the asymptotic MSE performance of the MT-MUSIC algorithm [16]. Based on this analysis, we obtain a strongly consistent estimator of the asymptotic MSE. The MSE estimator is then utilized to develop a data-driven procedure for optimal selection of the MT-function within some parametric family.

3.1. Asymptotic performance analysis

Throughout the analysis, the deterministic vector $\boldsymbol{\theta} \triangleq [\theta_1, ..., \theta_q]^T$ denotes the true DOAs. The random vector $\hat{\boldsymbol{\theta}}_u \triangleq [\hat{\theta}_1, ..., \hat{\theta}_q]^T$ will denote their estimates obtained by the MT-MUSIC. For the sake of simplicity, we shall assume that a sequence of i.i.d. samples from $P_{\mathbf{X}}$ is available and that the MT-mean $\boldsymbol{\mu}^{(u)}_{\mathbf{X}} = \mathbf{0}$. We note that the latter assumption holds whenever the MT-function $u(\cdot)$ is zero-centered and symmetric and observations are symmetrically distributed about the origin. Furthermore, it is assumed that condition A-2, stated above, is satisfied.

The following theorem states sufficient conditions for asymptotic normality and unbiasedness of the MT-MUSIC estimator.

Theorem 1. *Assume that the following conditions are satisfied:*
B-1) The expectation $\mathrm{E}[\|\mathbf{X}\|^4 u^2(\mathbf{X}); P_{\mathbf{X}}]$ *is finite.*
B-2) The steering vector $\mathbf{a}(\vartheta)$ *has a bounded Euclidean norm.*
B-3) $\mathbf{a}(\vartheta)$ *is twice continuously differentiable.*
B-4) The true DOAs $\theta_1, ..., \theta_q$ *lie in the interior of* Θ.
B-5) The eigenvalues $\lambda_1^{(u)}, ..., \lambda_q^{(u)}$ *associated with the signal subspace are distinct. Then,*

$$\sqrt{N}(\hat{\boldsymbol{\theta}}_u - \boldsymbol{\theta}) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{R}^{(u)}(\boldsymbol{\theta})) \quad as \quad N \to \infty, \tag{8}$$

where "$\xrightarrow{d}$*" denotes converges in distribution [25],*

$$\mathbf{R}^{(u)}(\boldsymbol{\theta}) \triangleq \mathrm{E}[\varphi_u^2(\mathbf{X})\boldsymbol{\Psi}_u(\mathbf{X}, \boldsymbol{\theta})\boldsymbol{\Psi}_u^T(\mathbf{X}, \boldsymbol{\theta}); P_{\mathbf{X}}], \tag{9}$$

$$[\boldsymbol{\Psi}_u(\mathbf{X}, \boldsymbol{\theta})]_i \triangleq \frac{\mathrm{Re}\{\dot{\mathbf{a}}^H(\theta_i)\mathbf{P}_{\mathbf{A}}^{\perp}\mathbf{X}\mathbf{X}^H\mathbf{E}^{(u)}\mathbf{a}(\theta_i)\}}{\dot{\mathbf{a}}^H(\theta_i)\mathbf{P}_{\mathbf{A}}^{\perp}\dot{\mathbf{a}}(\theta_i)}, \tag{10}$$

$$\mathbf{E}^{(u)} \triangleq \mathbf{A}(\mathbf{A}^H(\boldsymbol{\Sigma}_{\mathbf{X}}^{(u)} - \gamma^{(u)}\mathbf{I}_p)\mathbf{A})^{-1}\mathbf{A}^H, \tag{11}$$

$\varphi_u(\cdot)$ *is defined below (2),* $\dot{\mathbf{a}}(\vartheta) \triangleq d\mathbf{a}(\vartheta)/d\vartheta$, $[\mathbf{v}]_i$ *denotes the i-th coordinate of a vector* $\mathbf{v}$, $\mathbf{P}_{\mathbf{A}}^{\perp}$ *is the projection matrix onto the null-space of* $\mathbf{A}^H$, $\gamma^{(u)}$ *is defined in (6) and* $\mathbf{I}_p$ *denotes a* $p \times p$ *identity matrix.*

Theorem 1 implies that the asymptotic MSE matrix of the MT-MUSIC estimator is given by:

$$\mathbf{C}^{(u)}(\boldsymbol{\theta}) \triangleq N^{-1}\mathbf{R}^{(u)}(\boldsymbol{\theta}). \tag{12}$$

We note that asymptotic MSE (12) was obtained without restricting the probability distribution of the data, $P_{\mathbf{X}}$, to a specific family of probability distributions (e.g., Gaussian, elliptical). In particular, when the noise obeys a proper complex normal distribution and the MT-function $u(\cdot)$ is non-zero and constant valued, it can be shown that the asymptotic MSE (12) coincides with the one reported in [26, Eq. 3.11a], [27, Eq. 55] and [28, Eq. 69] for the standard SCM based MUSIC. This result is intuitive since for any non-zero and constant valued MT-function the MT-MUSIC coincides with the standard SCM based MUSIC.

In the following Proposition, a strongly consistent estimate of (12) is developed. This MSE estimator will be applied in the following subsection for optimal selection of the MT-function.

Proposition 1. *Define the empirical asymptotic MSE:*

$$\hat{\mathbf{C}}^{(u)}(\hat{\boldsymbol{\theta}}_u) \triangleq N^{-1}\hat{\mathbf{R}}^{(u)}(\hat{\boldsymbol{\theta}}_u), \tag{13}$$

where

$$\hat{\mathbf{R}}^{(u)}(\hat{\boldsymbol{\theta}}_u) \triangleq N\sum_{n=1}^{N}\hat{\varphi}_u^2(\mathbf{X}_n)\hat{\boldsymbol{\Psi}}_u(\mathbf{X}_n, \hat{\boldsymbol{\theta}}_u)\hat{\boldsymbol{\Psi}}_u^T(\mathbf{X}_n, \hat{\boldsymbol{\theta}}_u), \tag{14}$$

$$[\hat{\boldsymbol{\Psi}}_u(\mathbf{X}, \hat{\boldsymbol{\theta}}_u)]_i \triangleq \frac{\mathrm{Re}\{\dot{\mathbf{a}}^H(\hat{\theta}_i)\mathbf{P}_{\hat{\mathbf{A}}}^{\perp}\mathbf{X}\mathbf{X}^H\hat{\mathbf{E}}^{(u)}\mathbf{a}(\hat{\theta}_i)\}}{\dot{\mathbf{a}}^H(\hat{\theta}_i)\mathbf{P}_{\hat{\mathbf{A}}}^{\perp}\dot{\mathbf{a}}(\hat{\theta}_i)}, \tag{15}$$

$$\hat{\mathbf{E}}^{(u)} \triangleq \hat{\mathbf{A}}(\hat{\mathbf{A}}^H(\hat{\boldsymbol{\Sigma}}_{\mathbf{X}}^{(u)} - \hat{\gamma}^{(u)}\mathbf{I}_p)\hat{\mathbf{A}})^{-1}\hat{\mathbf{A}}^H, \tag{16}$$

$\hat{\varphi}_u(\cdot)$ *is defined below (4),* $\hat{\mathbf{A}} \triangleq [\mathbf{a}(\hat{\theta}_1), ..., \mathbf{a}(\hat{\theta}_q)]$, *and* $\hat{\gamma}^{(u)} \triangleq \mathrm{trace}\{\mathbf{P}_{\hat{\mathbf{A}}}^{\perp}\hat{\boldsymbol{\Sigma}}_{\mathbf{X}}^{(u)}\}/(p-q)$. *Furthermore, assume that conditions B-1−B-3, stated in Theorem 1, are satisfied. Then,*

$$N\|\hat{\mathbf{C}}^{(u)}(\hat{\boldsymbol{\theta}}_u) - \mathbf{C}^{(u)}(\boldsymbol{\theta})\| \xrightarrow{a.s.} 0 \quad as \quad N \to \infty, \tag{17}$$

where "$\xrightarrow{a.s.}$*" denotes almost sure (a.s.) convergence [25].*

3.2. Optimal selection of the MT-function

The MT-function $u(\cdot)$ is selected to minimize an empirical estimate of the asymptotic average RMSE. This estimate is obtained by taking the arithmetic mean over the square roots of the diagonal terms that comprise the empirical asymptotic MSE matrix (13). Hence, the empirical asymptotic average RMSE takes the form:

$$J_u \triangleq \frac{1}{q}\sum_{i=1}^{q}\frac{\sqrt{\sum_{n=1}^{N}\hat{\varphi}_u^2(\mathbf{X}_n)\mathrm{Re}^2\{\dot{\mathbf{a}}^H(\hat{\theta}_i)\mathbf{P}_{\hat{\mathbf{A}}}^{\perp}\mathbf{X}_n\mathbf{X}_n^H\hat{\mathbf{E}}^{(u)}\mathbf{a}(\hat{\theta}_i)\}}}{\dot{\mathbf{a}}^H(\hat{\theta}_i)\mathbf{P}_{\hat{\mathbf{A}}}^{\perp}\dot{\mathbf{a}}(\hat{\theta}_i)}. \tag{18}$$

We emphasize that the objective function (18) is constructed via the *same* sequence of data samples used for implementation of the MT-MUSIC algorithm. Here, the class of MT-functions is confined to some parametric family $\{u(\mathbf{X}; \boldsymbol{\omega}), \boldsymbol{\omega} \in \boldsymbol{\Omega} \subseteq \mathbb{C}^r\}$ that satisfies conditions A-1, A-2 and B-1−B-5 stated above. The optimal MT-function parameter $\boldsymbol{\omega}_{\text{opt}}$ is obtained via numerical minimization of the objective function $J_u(\boldsymbol{\omega})$. When $\boldsymbol{\omega}$ is one-dimensional, a simple line search can be implemented. Otherwise, the minimization can be carried out via greedy search or gradient descent [29].

Throughout the sequel of this paper, the MT-MUSIC algorithm implemented with MT-function whose parameters are selected according to the MSE-based optimization approach, discussed above, will be referred to as "optimized MT-MUSIC".

4. NUMERICAL EXAMPLES

In this section, we evaluate the performance of the optimized MT-MUSIC (MT-MUSIC$_{\text{opt}}$) as compared to the suboptimal MT-MUSIC (MT-MUSIC$_{\text{sub}}$) [16], the non-robust SCM based MUSIC (SCM-MUSIC) [1], [2] and to its robust extensons based on the empirical sign-covariance (SGN-MUSIC) [4] and Tyler's scatter M-estimator (TYLER-MUSIC) [5], [8]. The MT-MUSIC$_{\text{sub}}$ (implemented with a Gaussian MT-function) and TYLER-MUSIC are implemented exactly as described in the third and fourth paragraphs of Section VI in [16]. In all compared algorithms $K_{\Theta} = 10^5$ equally spaced samples of the parameter space Θ were used in order to obtain the empirical pseudo-spectra.

Here, the optimized MT-MUSIC is implemented with the following non-Gaussian parametric MT-function:

$$u(\mathbf{x}; \omega) \triangleq \|\mathbf{x}\|^{-\omega}, \quad \omega \in \mathbb{R}_+. \tag{19}$$

Notice that when $\omega = 0$ the resulting MT-covariance (3) coincides with the standard covariance. Furthermore, when $\omega = 2$, one can verify that for centered symmetric distributions, the MT-covariance is a scaled version of the sign-covariance [4]. Also note that for $\omega > 2$, the MT-function (19) satisfies the robustness conditions stated in (5) over a sufficiently large subset of $\mathbb{C}^p$ that does not contain the origin. To see this, define the set $\mathcal{R}_\epsilon \triangleq \{\mathbf{y} \in \mathbb{C}^p : \|\mathbf{y}\| > \epsilon\}$, where $\epsilon > 0$ is some fixed small positive constant. Clearly, $u(\mathbf{y}; \omega) < \epsilon^{-\omega}$ and $u(\mathbf{y}; \omega)\|\mathbf{y}\|^2 < \epsilon^{2-\omega}$ for any $\mathbf{y} \in \mathcal{R}_\epsilon$ and any fixed $\omega > 2$. Thus, since $P_{\mathbf{X}}(\mathcal{R}_\epsilon) \approx 1$ for sufficiently small ϵ we conclude that the empirical MT-covariance is B-robust with high probability, i.e., condition A-1 is satisfied with high probability. Furthermore, similarly to the proof of Proposition 4 in [16], it can be shown that, unlike the empirical sign-covariance [4] and Tyler's M-estimator of scatter [8], for any fixed $\omega > 2$ the influence function of the empirical MT-covariance [16, Eq. (17)], associated with the non-Gaussian MT-function (19), approaches zero as the outlier norm approaches infinity. This property results in enhanced robustness to large-norm outliers. Finally, it is important to note that the MT-function (19) is a spherically contoured and strictly-positive function. Hence, by [30, Th. 1], the key condition A-2, required for implementation of the MT-MUSIC, is satisfied when the noise obeys a CG distribution.

In the following simulation examples, the vector signal $\mathbf{S}_n$ in the array model (1) is comprised of $q = 2$ statistically independent BPSK signals with equal variance σ_S^2. The steering vector $\mathbf{a}(\theta) \triangleq \frac{1}{\sqrt{p}} \left[1, e^{-i\pi \sin(\theta)}, \ldots, e^{-i\pi(p-1)\sin(\theta)}\right]^T$ represents a uniform linear array with half wavelength spacing corresponding to a far-field narrow band signal with $p = 16$ elements. Here, the parameter space $\Theta = [-\pi/2, \pi/2)$. The DOAs were set to $\theta_1 = 0^\circ$ and $\theta_2 = 5^\circ$. We considered two types of zero-centered noise distributions with isotropic dispersion $\sigma_{\mathbf{Z}}^2 \mathbf{I}_p$: 1) Gaussian and 2) K-distributed CG noise [5] with shape parameter $\nu = 0.75$.

In the first simulation example, we compared the asymptotic average RMSE (obtained by taking the arithmetic mean over the square roots of the diagonal terms comprising (12)) to its empirical estimate (18) as a function of the power parameter ω of the MT-function (19) for sample size $N = 1000$. The generalized SNR (GSNR), defined here as $\text{GSNR} \triangleq 10 \log_{10} \sigma_S^2 / \sigma_{\mathbf{Z}}^2$, was set to -10 [dB]. Observing Figs. 1(a) (Gaussian noise) and 1(b) (K-distributed noise), one sees that due to the consistency of (13), that follows from Proposition 1, the compared quantities are close. This illustrates the reliability of (18) for optimal choice of the MT-function parameter, as discussed in subsection 3.2. By comparing Figs. 1(a) and 1(b), one can also notice that for the heavy-tailed K-distributed noise, the asymptotic average RMSE is minimized for a larger value of ω, that corresponds to a narrower MT-function, as compared to the light-tailed Gaussian noise. This result is intuitive since in the presence of heavy-tailed noise a narrower MT-function with faster decay toward zero should be applied to effectively mitigate the effect of outliers.

In the second simulation example, we compared the empirical average RMSE of the optimized MT-MUSIC to those obtained by the other compared estimators versus GSNR and samples size N. All empirical average RMSE curves were obtained via 10^4 Monte-Carlo trials. The optimal MT-function parameter ω_{opt} was obtained by minimizing (18) over $K_\Omega = 21$ equally spaced grid points of the interval $\Omega \triangleq [0, 10]$. For each type of comparison (GSNR, sample size), we also report the optimal asymptotic average RMSE of the MT-MUSIC for the considered MT-function (19). This quantity was obtained by minimizing the arithmetic mean of the square roots of the diagonal terms comprising (12) w.r.t. the power parameter $\omega \in \Omega$ in (19). The sample size in the GSNR analysis was set to $N = 1000$. In the sample size analysis, the GSNR was set to -11 [dB] and -21 [dB] for the Gaussian and K-distributed noise, respectively. Observing Figs. 1(c) and 1(e), one can notice that for the Gaussian noise, all compared estimators attain similar performance. Observing Figs. 1(d) and 1(f), one sees that for the heavy-tailed K-distributed noise, the optimized MT-MUSIC outperforms the non-robust SCM-MUSIC and all other robust alternatives. This performance advantage is a consequence of the following properties. First, as discussed above, the empirical MT-covariance associated with the MT-function (19) can gain enhanced robustness to large-norm outliers as compared to Tyler's and the sign-covariance estimators. Second, unlike all compared methods, the proposed approach involves optimization of a consistent estimate of the asymptotic average RMSE of the DOAs estimates.

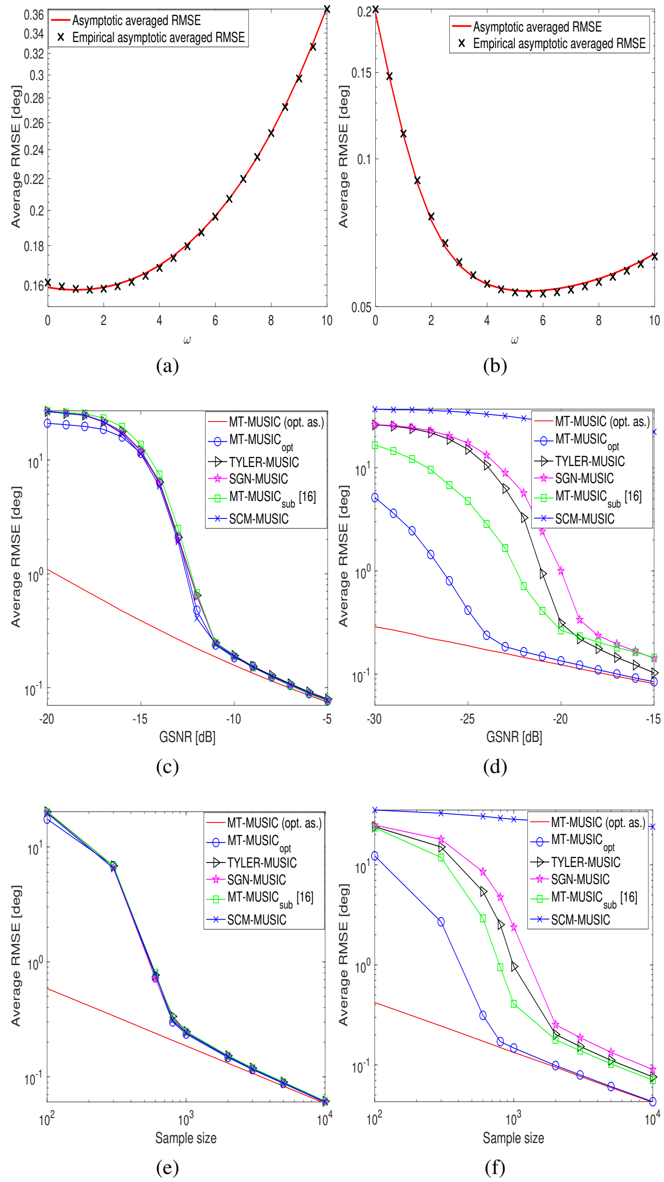

Fig. 1. **Estimation performance for Gaussian noise (left column) and K-distributed noise (right column)**. (a)+(b): Asymptotic average RMSE predicted by the theory and its empirical estimate (18) versus the power parameter ω of the non-Gaussian MT-function (19). (c)-(f): Optimum asymptotic (opt. as.) average RMSE of the MT-MUSIC and the empirical average RMSE of the MT-MUSIC$_{\text{opt}}$ as a function of GSNR (c)+(d) and sample size (e)+(f), as compared to the empirical average RMSEs of the other examined estimators.

5. CONCLUSION

In this paper we analyzed the asymptotic MSE performance of the MT-MUSIC algorithm [16]. Based on this analysis, a data-driven procedure for optimal selection of the MT-function parameters was developed that minimizes an empirical estimate of the asymptotic average RMSE. By specifying the MT-function in a family of non-Gaussian outlier-suppressing MT-functions, the optimized MT-MUSIC demonstrated significant performance advantage over the suboptimal MT-MUSIC [16] and other robust MUSIC extensions.

6. REFERENCES

[1] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation,* vol. 34, no. 3, pp. 276-280, 1986.

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

[3] T. W. Anderson, *An introduction to multivariate statistical analysis*. Wiley, 2003.

[4] S. Visuri, H. Oja and V. Koivunen, "Subspace-based direction-of-arrival estimation using nonparametric statistics," *IEEE Transactions on Signal Processing,* vol. 49, no. 9, pp. 2060-2073, 2001.

[5] E. Ollila, D. E. Tyler, V. Koivunen and H. V. Poor, "Complex elliptically symmetric distributions: survey, new results and applications," *IEEE Transactions on Signal Processing,* vol. 60, no. 1, pp. 5597-5625, 2012.

[6] R. A. Maronna, "Robust M-estimators of multivariate location and scatter," *The Annals of Statistics,* vol. 4, no. 1, pp. 51-67, 1976.

[7] P. J. Huber, *Robust statistics.* Wiley, 1981.

[8] D. E. Tyler, "A distribution-free M-estimator of multivariate scatter," *Annals of statistics,* vol. 15, no. 1, pp. 234-251, 1987.

[9] R. Couillet and A. Kammoun, "Robust G-MUSIC," *Proceedings of EUSIPCO 2014,* pp. 2155-2159.

[10] R. Couillet, "Robust spiked random matrices and a robust G-MUSIC estimator," *Journal of Multivariate Analysis,* vol. 140, pp. 139-161, 2015.

[11] P. Tsakalides and C. L. Nikias, "The robust covariation-based MUSIC (ROC-MUSIC) algorithm for bearing estimation in impulsive noise environments," *IEEE Transactions on Signal Processing,* vol. 44, no. 7, pp. 1623-1633, 1996.

[12] T. H. Liu and J. M. Mendel, "A subspace-based direction finding algorithm using fractional lower order statistics," *IEEE Transactions on Signal Processing,* vol. 49, no. 8, pp. 1605-1613, 2001.

[13] A. Swami and B. M. Sadler, "On some detection and estimation problems in heavy-tailed noise," *Signal Processing*, vol. 82, no. 12, pp. 1829–1846, 2002.

[14] C. H. Lim, S. Chong-Meng See, A. M. Zoubir and B. Poh Ng, "Robust adaptive trimming for high-resolution direction finding," *IEEE Signal Processing Letters,* vol. 16, no. 7, pp. 580-583, 2009.

[15] W. J. Zeng, H. C. So and L. Huang, "l_p-MUSIC: Robust direction-of-arrival estimator for impulsive noise environments," *IEEE Transactions on Signal Processing,* vol. 61, no, 17, pp. 4296-4308, 2013.

[16] K. Todros and A. O. Hero, "Robust multiple signal classification via probability measure transformation," *IEEE Transactions on Signal Processing*, vol. 63, no. 5, pp. 1156-1170, 2015.

[17] K. Todros and A. O. Hero, "On measure transformed canonical correlation analysis," *IEEE Transactions on Signal Processing*, vol. 60, no. 9, pp. 4570–4585, 2012.

[18] K. Todros and A. O. Hero, "Robust measure transformed MUSIC for DOA estimation," in *Proceedings of International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4190–4194, 2014.

[19] K. Todros and A. O. Hero, "Measure-transformed quasi-maximum likelihood estimation," *IEEE Transactions on Signal Processing*, vol. 65, no. 3, pp. 748–763, 2017.

[20] N. Halay, K. Todros and A. O. Hero, "Binary hypothesis testing via measure transformed quasi likelihood ratio test," *IEEE Transactions on Signal processing*, vol. 65, no. 24, pp. 6381-6396, Dec. 2017.

[21] N. Halay and K. Todros, "Plug-in measure-transformed quasi-likelihood ratio test for random signal detection," *IEEE Signal Processing Letters*, vol. 24, no. 6, pp. 838–842, 2017.

[22] K. Todros, "Robust composite binary hypothesis testing via measure transformed quasi score test," in *Signal Processing,* vol. 155, pp. 202-217, Feb. 2019.

[23] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw and W. A. Stahel, *Robust statistics: the approach based on influence functions*. John Wiley & Sons, 2011.

[24] G. B. Folland, *Real Analysis*. John Wiley and Sons, 1984.

[25] K. B. Athreya and S. N. Lahiri, *Measure theory and probability theory,* Springer-Verlag, 2006.

[26] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramér-Rao bound," *IEEE Transactions on Acoustics, Speech, and Signal Processing,* vol. 37, no. 5, pp. 720-741, 1989.

[27] J. F. Cardoso and E. Moulines, "Asymptotic performance analysis of direction-finding algorithms based on fourth-order cumulants," *IEEE Transactions on Signal Processing,* vol. 43, no. 1, pp. 214-224, 1995.

[28] J. P. Delmas, "Performance bounds and statistical analysis of DOA estimation," *Academic Press Library in Signal Processing,* vol. 3, 1st Edition, Array and Statistical Signal Processing, Elsevier, 2013.

[29] S. Boyd and and L. Vandenberghe, *Convex Optimization.* Cambridge Univ. Press, 2004.

[30] N. Halay and K. Todros, "A note on the measure-transformed MUSIC algorithm," *Technical report.* Online version: `http://www.ee.bgu.ac.il/~todros/Comment.pdf`