

Publication IV

Esa Ollila and Visa Koivunen. 2009. Influence function and asymptotic efficiency of scatter matrix based array processors: case MVDR beamformer. IEEE Transactions on Signal Processing, volume 57, number 1, pages 247-259.

© 2008 Institute of Electrical and Electronics Engineers (IEEE)

Reprinted, with permission, from IEEE.

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of Aalto University School of Science and Technology's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org.

By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

Influence Function and Asymptotic Efficiency of Scatter Matrix Based Array Processors: Case MVDR Beamformer

Esa Ollila, *Member, IEEE*, and Visa Koivunen, *Senior Member, IEEE*

Abstract—In this paper, we consider array processors that are scale-invariant functions of the array covariance matrix. The emphasis is on Capon’s MVDR beamformer. We call such an array processor as scatter matrix based (SMB) array processor since the covariance matrix is required only up to a constant scalar and thus a scatter matrix (proportional to covariance under finite covariance assumption) provides sufficient information. In order to establish interesting statistical robustness and large sample properties, we derive a general expression for the influence function and the asymptotic covariance structure of SMB-MVDR beamformer weights. Our results apply under the class of complex elliptically symmetric distributions, which includes the commonly used complex normal distribution as a special case. We illustrate the theory by deriving the IF and asymptotic relative efficiencies of the conventional SMB-MVDR beamformer that employs the sample covariance matrix and beamformers that employ robust M -estimators of scatter. Theoretical findings are confirmed by simulations. Our findings favor beamformers based upon M -estimators of scatter, since they combine a high efficiency with appealing robustness properties.

Index Terms— M -estimation, beamforming, complex elliptical distributions, influence function, robustness, statistical efficiency, statistical functional.

I. INTRODUCTION

OPTIMAL array processors are derived under idealized assumptions, e.g., that the covariance matrix $\mathbf{C}(\mathbf{z}) = E[\mathbf{z}\mathbf{z}^H]$ of the array output vector $\mathbf{z} \in \mathbb{C}^k$ is known, and some times, the distribution F of \mathbf{z} is assumed to be known also, e.g., the conventional (circular) complex normal (CN) distribution. The resulting array processor is often found to be a function of the covariance matrix alone, i.e., of the form $\mathbf{h}(\mathbf{C})$, where $\mathbf{h} : \text{PDH}(k) \rightarrow \mathbb{C}^q$ is a function that maps the covariance matrix $\mathbf{C}(\mathbf{z}) \in \text{PDH}(k)$ to a vector in complex Euclidean q -space \mathbb{C}^q , where $\text{PDH}(k)$ denotes the set of complex positive definite

Hermitian $k \times k$ matrices. If the array processor \mathbf{h} is *scale-invariant* (positively homogeneous function of degree zero), i.e.,

$$\mathbf{h}(\mathbf{C}) = \mathbf{h}(a\mathbf{C}) \text{ for all } a > 0 \text{ (and } \mathbf{C} \in \text{PDH}(k)) \quad (1)$$

then it requires covariance matrix \mathbf{C} only up to a constant scalar, and thus requires information on the “scatter” of the data set only but not on the “scale.” Hence, one may replace $\mathbf{C}(\mathbf{z})$ in $\mathbf{h}(\mathbf{C})$ by a more general notion of covariance, called the *scatter matrix* $\mathbf{C}(\mathbf{z})$, which is proportional to $\mathbf{C}(\mathbf{z})$ under finite variance assumption under the class of circular complex elliptically symmetric (CES) [1] distributions that includes the widely used CN distribution as an important special case. Covariance matrix is an example of a scatter matrix but more general scatter matrices exists which do not rely upon finiteness of variances, e.g., robust M -estimators of scatter [2], [3]. We call resulting array processor $\mathbf{h}(\mathbf{C})$ as a *scatter matrix based (SMB) array processor*.

For example, the classical Capon’s [4] *minimum variance distortionless response (MVDR)* beamformer weight vector

$$\mathbf{w}_{\mathbf{C}}(\mathbf{z}) = \frac{\mathbf{C}(\mathbf{z})^{-1}\mathbf{g}}{\mathbf{g}^H\mathbf{C}(\mathbf{z})^{-1}\mathbf{g}}, \quad (2)$$

where \mathbf{g} is the nominal *array response vector* (assumed to be known exactly), satisfies criteria (1). In this paper, we define and study a wide class of MVDR beamformers by replacing covariance matrix in (2) by a scatter matrix $\mathbf{C}(\mathbf{z})$, yielding a SMB-MVDR beamformer weight vector $\mathbf{w}_{\mathbf{C}}(\mathbf{z})$. In practice, since the employed scatter matrix is unknown, an intuitive approach is to replace the true scatter matrix by an estimated value $\hat{\mathbf{C}}$ which yields the *estimated SMB-MVDR beamformer* $\hat{\mathbf{w}}_{\mathbf{C}}$. As an example, if the covariance matrix is our choice for the scatter matrix, then the commonly used estimate is the *sample covariance matrix (SCM)* $\hat{\mathbf{C}} = 1/n \sum_{i=1}^n \mathbf{z}_i\mathbf{z}_i^H$ since it is the maximum-likelihood estimator (MLE) under the conventional assumption that the array output data $\mathbf{z}_1, \dots, \mathbf{z}_n$ is a random sample from CN distribution.

The estimated array processor $\mathbf{h}(\hat{\mathbf{C}})$ employing the SCM may yield optimal estimator if the assumptions under which the theoretical array processor \mathbf{h} was derived holds. However, it may have poor performance when the nominal assumptions are not valid, e.g., in the face of outliers (outliers occur in array data e.g., due to heavy-tailed, impulsive noise such as man-made interference) or (slight/large) departures from nominal distributional assumptions. Hence, it may be advisable to employ some other estimator of covariance (instead of SCM) that possess better robustness and statistical efficiency characteristics.

Manuscript received February 06, 2007; revised August 27, 2008. First published October 31, 2008; current version published January 06, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. A. Rahim Leyman.

E. Ollila is with the Department of Mathematical Sciences, University of Oulu, FIN-90014 Oulu, Finland, and also with the Signal Processing Laboratory, SMARAD CoE, Helsinki University of Technology, Espoo, FIN-02015 HUT (e-mail: esa.ollila@oulu.fi).

V. Koivunen is with the Signal Processing Laboratory, SMARAD CoE, Helsinki University of Technology, Espoo, FIN-02015 HUT (e-mail: visa@wooster.hut.fi).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSP.2008.2007347

In this paper, we focus on MVDR beamformer and address the following questions:

- i) How *robust* to outliers is the MVDR beamformer employing an estimated covariance matrix?
- ii) What is the *efficiency* (performance loss, gain) of the MVDR beamformer employing the SCM as opposed to some other estimator of covariance when the assumption of normality is not valid?

We address the questions i) and ii) via *asymptotic (large sample) analysis* of $\hat{\mathbf{w}}_C$ under the assumption that $\mathbf{z}_1, \dots, \mathbf{z}_n$ is a random sample from a CES distribution. To be more specific, general expressions for the influence function (IF) [5] and the asymptotic covariance matrix of the SMB-MVDR weight vector are established. Our investigations reveal that a single, easily computable scalar-index of asymptotic relative efficiency (ARE), defined as a ratio between matrix traces of the asymptotic covariance matrices, is sufficient and that the robustness (boundedness and continuity) of the influence function can be described by a real valued scalar function. Our findings thus substantially facilitate robustness and accuracy (efficiency) comparisons of different SMB-MVDR beamformers. We illustrate the general theory by calculating the IFs and AREs of the SMB-MVDR weight vectors employing SCM and selected complex M -estimators of scatter. Theoretical findings are confirmed via simulation studies comparing empirical influence function and finite sample relative efficiencies with the obtained theoretical counterparts. Our findings favor SMB-MVDR beamformer based upon robust M -estimators, since they combine a high efficiency with appealing robustness properties.

We wish to point out that in beamforming literature, “robust” more commonly refers to *robustness to steering errors* (imprecise knowledge of the array response \mathbf{g} may be due to uncertainty in array element locations, steering directions and calibration errors) and *robustness in the face of insufficient sample support* that may lead to rank deficient SCM or inaccurate estimates of the array covariance matrix. The diagonal loading of the SCM is one of the most popular techniques to overcome the above problems, i.e., to use $(\hat{\mathbf{C}} + \gamma \mathbf{I})$, $\gamma \in \mathbb{R}$, in place of $\hat{\mathbf{C}}$, which may not be full rank and hence not invertible. For this type of robustness studies, see e.g., [6]–[9] and references therein. In this paper, the term “robust” refers to statistical robustness to outliers [5], commonly measured by the concept of the influence function. We wish to point out that robustness (as measured by the influence function) of the MVDR beamformer remains unaltered by diagonally loading the covariance matrix \mathbf{C} , i.e., using $\mathbf{C}_\gamma(F) = \mathbf{C}(F) + \gamma \mathbf{I}$, where γ is some constant diagonal loading term not dependent on the distribution F of \mathbf{z} . Although (statistical) robustness of the MVDR weight functional is not improved with diagonal loading, it provides, however, other kind of robustness by improving the condition number of the estimated array covariance matrix. Naturally, IF is an asymptotic concept, and it is not the correct tool to analyze the performance in sample starved scenarios.

The paper is organized as follows. Section II reviews the family of CES distributions. In Section III, scatter matrix and the SMB-MVDR beamformer class are defined formally. Also robust M -estimators of scatter matrix are reviewed. In

Section IV, our tools, the IF and asymptotic relative efficiency, are discussed. It is highlighted that establishing the asymptotic covariance structure of a statistic requires calculation of (asymptotic) covariance matrix and pseudo-covariance matrix. The statistical robustness and efficiency analysis begins in Section V, where the IF and asymptotic covariance structure of SMB-MVDR beamformers are derived. The efficiencies and robustness of SMB-MVDR beamformers based on SCM and selected M -estimators of scatter are compared as well. In Section VI, the simulated empirical influence function and finite sample relative efficiencies of the SMB-MVDR beamformers neatly validate the theoretical (asymptotic) findings. Section VII concludes, and the Appendix presents all the proofs.

Notation: Superscripts $(\cdot)^H$, $(\cdot)^T$, and $(\cdot)^*$ stand for the Hermitian transpose, transpose and complex conjugate, respectively. Symbol j denotes the imaginary unit, $|\cdot|$ denotes the matrix determinant (or, complex modulus when its argument is a complex scalar), $\text{Tr}(\cdot)$ denotes the matrix trace, $\text{Re}(\cdot)$ extracts the real part of its argument and $\|\cdot\|$ denotes the norm of a vector (i.e., $\|\mathbf{z}\| = \sqrt{\mathbf{z}^H \mathbf{z}}$). Furthermore, symbol $\stackrel{d}{\sim}$ reads “has the same distribution as” and \xrightarrow{L} means convergence in distribution or in law. By convention, $\mathbf{T}(F) \equiv \mathbf{T}(\mathbf{z})$ whenever $\mathbf{z} \sim F$ and for any statistical functional \mathbf{T} , e.g., $\mathcal{C}(F) \equiv \mathcal{C}(\mathbf{z})$.

II. COMPLEX ELLIPTICALLY SYMMETRIC DISTRIBUTIONS

Complex random vector (r.v.) \mathbf{z} has *circularly symmetric (CS) distribution* if $\mathbf{z} \stackrel{d}{=} \exp(j\theta)\mathbf{z}$ for all $\theta \in \mathbb{R}$. A prominent example in this class is the CN distribution: A zero mean random vector (r.v.) $\mathbf{z} = \mathbf{x} + j\mathbf{y} \in \mathbb{C}^k$ has *k -variate complex circular normal distribution*, labeled Φ_k , if \mathbf{x} and \mathbf{y} have joint $2k$ -variate real normal distribution and a $2k \times 2k$ real covariance matrix of a special form

$$E[\mathbf{x}\mathbf{x}^T] = E[\mathbf{y}\mathbf{y}^T] \quad \text{and} \quad E[\mathbf{x}\mathbf{y}^T] = -E[\mathbf{y}\mathbf{x}^T]. \quad (3)$$

The CN distribution has density [10]–[12] $f_{\mathbf{C}}(\mathbf{z}) = \pi^{-k} |\mathbf{C}|^{-1} \exp(-\mathbf{z}^H \mathbf{C}^{-1} \mathbf{z})$ if the covariance matrix \mathbf{C} is nonsingular. If \mathbf{C} is singular, then CN distribution do not have density, but the characteristic function (cf.) $\phi_{\mathbf{C}}(\mathbf{z}) = \exp(-(1/2)\mathbf{z}^H \mathbf{C} \mathbf{z})$ always exists and is unique. We shall write $\mathbf{z} \sim \mathcal{CN}_k(\mathbf{C})$. A natural extension of CN distribution is the class of (circular) *complex elliptically symmetric (CES) distributions* studied in [1]; see also [3] and [13].

Definition 1: Random vector $\mathbf{z} \in \mathbb{C}^k$ is said to have a (centered) CES distribution with parameter $\Sigma \in \text{PDH}(k)$, if its p.d.f. is of the form

$$f_{\Sigma}(\mathbf{z}) = c_{k,g} |\Sigma|^{-1} g(\mathbf{z}^H \Sigma^{-1} \mathbf{z}) \quad (4)$$

where $g : [0, \infty) \rightarrow [0, \infty)$ is a fixed function, called the *density generator*, independent of Σ and $c_{k,g}$ is a normalizing constant. We shall write $\mathbf{z} \sim F_{\Sigma} \equiv \mathcal{CE}_k(\Sigma, g)$.

In (4), $c_{k,g}$ is defined as $c_{k,g} = 2(s_k \mu_{k,g})^{-1}$, where $s_k = 2\pi^k / \Gamma(k)$ is the surface area of unit complex k -sphere $\mathbb{C}S^k = \{\mathbf{z} \in \mathbb{C}^k : \|\mathbf{z}\| = 1\}$ and $\mu_{k,g} = \int_0^\infty t^{k-1} g(t) dt$. Naturally, $c_{k,g}$ could be absorbed into the function g , but with this notation g can be independent of the dimension k . Observe that the regions of constant contours are ellipsoids in complex Euclidean

k -space, thus explaining the name for this class of distributions. CES distributions can also be defined more generally (without the existence of a density) via their characteristic function [1]. A generalization of CES distributions (avoiding the circularity assumption) is given in [13].

The functional form of the density generator $g(\cdot)$ uniquely distinguishes different CES distributions from another. For example, $g(t) = \exp(-t)$ yields the CN distribution, and

$$g(t) = \left(1 + \frac{2t}{\nu}\right)^{-(2k+\nu)/2}$$

yields the k -variate *complex t -distribution* [1], [3] with ν degrees of freedom ($0 < \nu < \infty$), labeled $T_{k,\nu}$. The case $\nu = 1$ is called the *complex Cauchy distribution*, and the limiting case $\nu \rightarrow \infty$ yields the CN distribution.

When the covariance matrix of F_{Σ} exists it is proportional to Σ , namely

$$\mathcal{C}(F_{\Sigma}) = \sigma_{\mathcal{C}} \Sigma \quad (5)$$

where $\sigma_{\mathcal{C}} = \sigma_{\mathcal{C}}(t) = k^{-1}E(t)$ and the *quadratic form* $t = \mathbf{z}^H \Sigma^{-1} \mathbf{z}$ is a positive real random variable (r.va.) with a p.d.f.

$$q(t) = t^{k-1} g(t) \mu_{k,g}^{-1}. \quad (6)$$

Hence, the covariance matrix of F_{Σ} exists if and only if $E(t) < \infty$, i.e., $\mu_{k+1,g} = \int t^k g(t) dt < \infty$. At Φ_k -distribution, $\sigma_{\mathcal{C}} = 1$, so $\mathcal{C}(\Phi_k) = \Sigma$ as expected. At $T_{k,\nu}$ -distribution, $\sigma_{\mathcal{C}} = \nu/(\nu - 2)$, indicating that $T_{k,\nu}$ do not possess a finite covariance matrix for $\nu \leq 2$.

III. SCATTER MATRIX BASED MVDR BEAMFORMER

A. Scatter Matrix

Definition 2: Let $\mathbf{s} = \mathbf{A}\mathbf{z}$ denote invertible linear transformation of $\mathbf{z} \in \mathbb{C}^k$ for any nonsingular $\mathbf{A} \in \mathbb{C}^{k \times k}$. Functional $\mathbf{C}(F) \in \text{PDH}(k)$ is called a *scatter matrix* if $\mathbf{C}(\mathbf{s}) = \mathbf{A}\mathbf{C}(\mathbf{z})\mathbf{A}^H$.

Equivariance property implies that

$$\mathbf{C}(F_{\Sigma}) = \sigma_{\mathcal{C}} \Sigma \quad (7)$$

for some positive *scalar factor* $\sigma_{\mathcal{C}} = \sigma_{\mathcal{C}}(t)$, the value of which depends both on the functional \mathbf{C} and on the underlying CES distribution F_{Σ} of \mathbf{z} only through the density (6) of the quadratic form $t = \mathbf{z}^H \Sigma^{-1} \mathbf{z}$. Equations (5) and (7) show that scatter matrix \mathbf{C} (provided it exists) is proportional to covariance matrix \mathcal{C} at all CES distributions F_{Σ} with finite covariance matrix. Thus, scatter matrix can be referred to as *generalized covariance matrix*, as it is more general concept: $\mathbf{C}(F_{\Sigma})$ can exist for CES distributions F_{Σ} which do not have finite covariance. Classical example of a scatter matrix is the covariance matrix $\mathcal{C}(\mathbf{z})$.

Let F_n denote the empirical distribution function associated with the data set $Z_n = \{\mathbf{z}_1, \dots, \mathbf{z}_n\}$. Then a natural *plug-in* estimator of $\mathbf{C}(F)$ is $\hat{\mathbf{C}} \equiv \mathbf{C}(F_n)$. For example, if $\mathbf{C}(F) = \mathcal{C}(F)$, then the (plug-in) estimator is the SCM $\hat{\mathbf{C}}$ since $\mathcal{C}(F_n) = \int \mathbf{z}\mathbf{z}^H dF_n(\mathbf{z}) = \hat{\mathbf{C}}$. At the finite sample level, equivariance under linear transformations implies that for any

nonsingular $\mathbf{B} \in \mathbb{C}^{k \times k}$, the estimator for the transformed data set $\mathbf{B}\mathbf{z}_1, \dots, \mathbf{B}\mathbf{z}_n$ is $\mathbf{B}\hat{\mathbf{C}}\mathbf{B}^H$.

B. M -Estimators of Scatter

M -estimators of multivariate scatter were first introduced in [2] for real data and later generalized in [3], [14], and [15] for complex data. As in the real case, they can be defined by generalizing MLE.

Let $\mathbf{z}_1, \dots, \mathbf{z}_n$ be an i.i.d. sample from a CES distribution F_{Σ} , where $n > k$ (i.e., sample size n is larger than the number of sensors k). The MLE of Σ , is found by minimizing the negative of the log-likelihood function, $l(\Sigma) = n \log |\Sigma| - \sum_{i=1}^n \log g(\mathbf{z}_i^H \Sigma^{-1} \mathbf{z}_i)$. By differentiating $l(\Sigma)$ with respect to Σ (using complex matrix differentiation rules [16]), shows that the MLE is a solution to estimating equation

$$\Sigma = \frac{1}{n} \sum_{i=1}^n \varphi_{\text{ml}}(\mathbf{z}_i^H \Sigma^{-1} \mathbf{z}_i) \mathbf{z}_i \mathbf{z}_i^H \quad (8)$$

where $\varphi_{\text{ml}}(x) = -g'(x)/g(x)$ is a weight function. For the CN distribution [i.e., $g(t) = \exp(-t)$], we have that $\varphi_{\text{ml}} \equiv 1$, which yields the SCM as the MLE of Σ . The MLE for $T_{k,\nu}$ distribution, labeled $\text{MLT}(\nu)$, is obtained with

$$\varphi_{\text{ml}}(x) = \frac{2k + \nu}{\nu + 2x}. \quad (9)$$

Note that $\text{MLT}(1)$ is the highly robust estimator corresponding to MLE of Σ for the complex Cauchy distribution.

We generalize (8), by defining *M -estimator of scatter*, denoted by $\hat{\mathbf{C}}_{\varphi}$, as the choice of $\mathbf{C} \in \text{PDH}(k)$ that solves the estimating equation

$$\mathbf{C} = \frac{1}{n} \sum_{i=1}^n \varphi(\mathbf{z}_i^H \mathbf{C}^{-1} \mathbf{z}_i) \mathbf{z}_i \mathbf{z}_i^H \quad (10)$$

where φ is a real-valued *weight function* on $[0, \infty)$. Hence, M -estimators is a wide class that include the MLEs for CES distributions as important special cases. M -estimators can be calculated by a simple iterative algorithm described in the Appendix A. The theoretical (population) counterpart, the *M -functional of scatter*, denoted by $\mathbf{C}_{\varphi}(F)$ or $\mathbf{C}_{\varphi}(\mathbf{z})$, is defined analogously as the solution of

$$\mathbf{C}(F) = E[\varphi(\mathbf{z}^H \mathbf{C}(F)^{-1} \mathbf{z}) \mathbf{z}\mathbf{z}^H]. \quad (11)$$

Observe that (11) reduces to (10) when F is the empirical distribution F_n , i.e., the solution $\hat{\mathbf{C}}_{\varphi}$ of (10) is the plug-in estimator $\mathbf{C}_{\varphi}(F_n)$ of \mathbf{C}_{φ} . It is easy to show that M -functional of scatter is equivariant in the sense of Definition 2. Due to equivariance, $\mathbf{C}_{\varphi}(F_{\Sigma}) = \sigma_{\varphi} \Sigma$, where the scalar factor $\sigma_{\varphi} = \sigma_{\varphi}(t)$ may be found by solving

$$E \left[\frac{\varphi\left(\frac{t}{\sigma_{\varphi}}\right) t}{\sigma_{\varphi}} \right] = k \quad (12)$$

where t has density (6). Often σ_{φ} need to be solved numerically from (12) but in some cases an analytic expression can be derived.

A prominent robust M -estimator, the Huber's M -estimator, labeled HUB(q), is defined by

$$\varphi(x) = \begin{cases} \frac{1}{b}, & \text{for } x \leq c^2 \\ \frac{c^2}{(xb)}, & \text{for } x > c^2 \end{cases}$$

where c is a tuning constant defined so that $q = F_{\chi_{2k}^2}(2c^2)$ for a chosen q ($0 < q \leq 1$) and the scaling factor $b = F_{\chi_{2(k+1)}^2}(2c^2) + c^2(1-q)/k$, where $F_{\chi_k^2}$ denotes the c.d.f. of chi-squared distribution with k degrees of freedom.¹ Note that the choice $q = 1$ yields $\varphi \equiv 1$, and thus HUB (1) correspond to the SCM.

By (10), $\hat{\mathbf{C}}_\varphi$ can be interpreted as a weighted covariance matrix. Hence, a robust weight function $\varphi(\cdot)$ should descend to zero. This means that small weights are given to those observations \mathbf{z}_i that are highly *outlying* in terms of measure $\mathbf{z}_i^H \hat{\mathbf{C}}_\varphi^{-1} \mathbf{z}_i$. Note that SCM is an M -estimator that gives unit weight ($\varphi \equiv 1$) to all observations. Fig. 1(a) plots the weight function (9) of MLT(ν) estimators for selected values of ν . Note that weight function (9) tends to weight function $\varphi \equiv 1$ of the SCM as expected (since $T_{k,\nu}$ tends to Φ_k distribution when $\nu \rightarrow \infty$). Thus, $\text{MLT}(\nu) \approx \hat{\mathbf{C}}$ for large values of ν . Fig. 1(b) depicts weight function of HUB(q) estimators for selected values of q . There is a tradeoff between robustness and efficiency: low values of q increase robustness (i.e., they give lower weight) but decrease efficiency at the nominal CN model. Naturally, $\text{HUB}(q) \approx \hat{\mathbf{C}}$ for $q \approx 1$.

C. SMB-MVDR Beamformer

The MVDR beamformer chooses \mathbf{w} as the minimizer of the output power $E(|y|^2) = \mathbf{w}^H \mathbf{C}(\mathbf{z}) \mathbf{w}$ while constraining the *beam response* along a *specific* look direction θ_0 and frequency ω_0 to be unity:

$$\min_{\mathbf{w}} \mathbf{w}^H \mathbf{C}(\mathbf{z}) \mathbf{w} \text{ subject to } \mathbf{w}^H \mathbf{g} = 1, \text{ where } \mathbf{g} = \mathbf{g}(\theta_0, \omega_0)$$

and $\mathbf{g}(\theta, \omega)$ denotes the *array transfer function* (array response, steering vector) whose functional form is assumed to be known, i.e., array is calibrated. Note that in narrowband applications the dependency on ω can be dropped. The solution to this constrained optimization problem is given by $\mathbf{w}_{\mathbf{C}}(\mathbf{z})$ in (2). Note that MVDR beamformer do not make any assumption on the structure of covariance matrix and hence can be considered as a "nonparametric method" [17]. The SMB-MVDR beamformer is defined formally below.

Definition 3: Scatter matrix based MVDR (SMB-MVDR) weight is defined as

$$\mathbf{w}_{\mathbf{C}}(\mathbf{z}) = \frac{\mathbf{C}(\mathbf{z})^{-1} \mathbf{g}}{\mathbf{g}^H \mathbf{C}(\mathbf{z})^{-1} \mathbf{g}}$$

where $\mathbf{C}(\mathbf{z})$ is a scatter matrix.

Define

$$\mathbf{w} = \mathbf{W} \mathbf{g}, \quad \text{where } \mathbf{W} = \frac{\Sigma^{-1}}{\mathbf{g}^H \Sigma^{-1} \mathbf{g}}. \quad (13)$$

¹The scaling factor b is chosen so that $E[\varphi(t)t] = k$ when $\mathbf{z} \sim \Phi_k$. This choice assures by (12), that $\sigma_\varphi = 1$ and thus $\mathbf{C}_\varphi(\Phi_k) = \Sigma$.

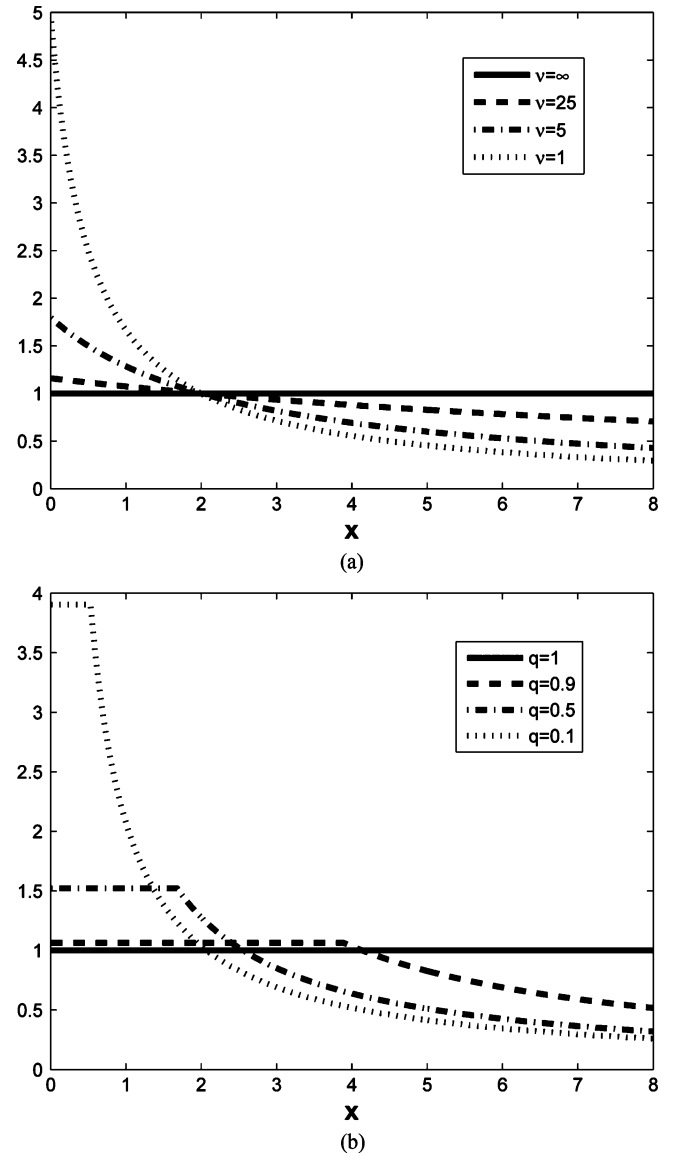


Fig. 1. (a) $\varphi(x)$ of MLT(ν) estimators. (b) $\varphi(x)$ of HUB(q) estimators.

Since the scatter matrix $\mathbf{C}(\mathbf{z})$ is proportional to Σ [cf. equation (7)] at F_{Σ} , it follows that

$$\mathbf{w}_{\mathbf{C}}(F_{\Sigma}) = \mathbf{w}$$

provided that $\mathbf{C}(F_{\Sigma})$ exists. For example, $\mathbf{w}_{\mathbf{C}}(F_{\Sigma}) = \mathbf{w}$ for all CES distributions F_{Σ} possessing finite covariance. Since in practice, the true scatter matrix $\mathbf{C}(F)$ is unknown, we replace it by the plug-in estimator $\hat{\mathbf{C}} = \mathbf{C}(F_n)$, which yields $\hat{\mathbf{w}}_{\mathbf{C}} = \mathbf{w}_{\mathbf{C}}(F_n)$ as a plug-in estimator for the SMB-MVDR weight.

Notation: SMB-MVDR weight functional (respectively, estimator) based on M -functional of scatter \mathbf{C}_φ is denoted by \mathbf{w}_φ (respectively, $\hat{\mathbf{w}}_\varphi$). The optimal weight functional (respectively, estimator) at F_{Σ} employs MLE of Σ (i.e., $\varphi = \varphi_{\text{ml}}$) and is denoted by \mathbf{w}_{mle} (respectively, $\hat{\mathbf{w}}_{\text{mle}}$).

IV. TOOLS TO COMPARE ESTIMATORS

A. Influence Function

Main tool for our statistical analysis of an estimator is the concept of influence function. It is a versatile tool for studying qualitative robustness (local stability) and large sample properties of estimators; see [5] and [18]. Denote the complex point mass distribution function at \mathbf{t} by $\Delta_{\mathbf{t}}(\mathbf{z})$ and consider the ε -contaminated distribution $F_{\varepsilon, \mathbf{t}}(\mathbf{z}) = (1 - \varepsilon)F(\mathbf{z}) + \varepsilon\Delta_{\mathbf{t}}(\mathbf{z})$. Then the influence function of a statistical functional \mathbf{T} at the distribution F is

$$\text{IF}(\mathbf{t}; \mathbf{T}, F) = \lim_{\varepsilon \downarrow 0} \frac{\mathbf{T}(F_{\varepsilon, \mathbf{t}}) - \mathbf{T}(F)}{\varepsilon} = \left. \frac{\partial}{\partial \varepsilon} \mathbf{T}(F_{\varepsilon, \mathbf{t}}) \right|_{\varepsilon=0}. \quad (14)$$

One may interpret the influence function as describing the effect (influence) of an *infinitesimal* point-mass contamination at \mathbf{t} on the estimator, standardized by the mass of the contamination. Hence, the IF gives asymptotic bias caused by the contamination. See [5] for a more detailed explanation of the influence function.

A robust estimator $\hat{\boldsymbol{\theta}} = \mathbf{T}(F_n)$ should have a bounded and continuous IF. Loosely speaking, the boundedness implies that a small amount of contamination at any point \mathbf{t} does not have an arbitrarily large influence on the estimator whereas the continuity implies that the small changes in the data set cause only small changes in the estimator. Note however that the IF is an *asymptotic concept*, characterizing stability of the estimator as n approaches infinity. Corresponding finite sample version is obtained by suppressing the limit in (14) and choosing $\varepsilon = 1/(n+1)$ and $F = F_n$. This yields the *empirical influence function* (EIF) (also called sensitivity function [5]) of the estimator $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}(Z_n)$:

$$\text{EIF}(\mathbf{t}; \hat{\boldsymbol{\theta}}, Z_n) = (n+1)[\hat{\boldsymbol{\theta}}(Z_n \cup \{\mathbf{t}\}) - \hat{\boldsymbol{\theta}}(Z_n)].$$

The EIF thus calculates the standardized effect of an additional observation at \mathbf{t} on the estimator. In many cases, $\text{EIF}(\mathbf{t}; \hat{\boldsymbol{\theta}}, Z_n)$ is a consistent estimator of $\text{IF}(\mathbf{t}; \mathbf{T}, F)$ [5], [19].

A fundamental result on the form of IF of *any* scatter matrix functional at CES distribution F_{Σ} was found in [14] and is stated below for further reference. The result is analogous to the corresponding result in the real case [5, p. 276], [20, Lemma 1].

Lemma 1: For any scatter matrix functional \mathbf{C} possessing an influence function, there exists two functions $\alpha_{\mathbf{C}}, \beta_{\mathbf{C}} : [0, \infty) \rightarrow \mathbb{R}$ such that

$$\text{IF}(\mathbf{t}; \mathbf{C}, F_{\Sigma}) = \alpha_{\mathbf{C}}(r^2)\mathbf{A}(\mathbf{u}\mathbf{u}^H - k^{-1}\mathbf{I})\mathbf{A}^H + \beta_{\mathbf{C}}(r^2)\boldsymbol{\Sigma}$$

where $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^H$ (nonsingular $\mathbf{A} \in \mathbb{C}^{k \times k}$), $r^2 = \|\mathbf{A}^{-1}\mathbf{t}\|^2 = \mathbf{t}^H \boldsymbol{\Sigma}^{-1} \mathbf{t}$ and $\mathbf{u} = \mathbf{A}^{-1}\mathbf{t}/r$.

Lemma 1 implies that the IF of \mathbf{C} is bounded if and only if the corresponding ‘‘weight functions’’ $\alpha_{\mathbf{C}}(\cdot)$ and $\beta_{\mathbf{C}}(\cdot)$ are bounded.

B. Asymptotic Performance of an Estimator

If r.v. \mathbf{z} satisfies (3), then it is said to be *second-order circular* [21] (or, proper [22]). Asymptotic performance analysis

of an estimator calls for a more general notion of complex normality that avoids the unnecessary second-order circularity assumption: r.v. $\mathbf{z} = \mathbf{x} + j\mathbf{y}$ is said to have a *generalized complex normal (GCN) distribution* if \mathbf{x} and \mathbf{y} have joint $2k$ -variate real normal distribution. The cf. of GCN distribution is [21], [23]

$$\phi_{\mathbf{C}, \mathcal{P}}(\mathbf{z}) = \exp \left[-\frac{1}{4} \{ \mathbf{z}^H \mathbf{C} \mathbf{z} + \text{Re}(\mathbf{z}^H \mathcal{P} \mathbf{z}^*) \} \right]$$

where $\mathcal{P} = \mathcal{P}(\mathbf{z}) = E[\mathbf{z}\mathbf{z}^T]$ is called the *pseudo-covariance matrix*. We shall write $\mathbf{z} \sim \text{CN}_k(\mathbf{C}, \mathcal{P})$. As expected, cf. of GCN distribution reduce to cf. of CN distribution when $\mathcal{P} = \mathbf{0}$ [which is the complex form of (3)], i.e., $\text{CN}_k(\mathbf{C}) = \text{CN}_k(\mathbf{C}, \mathbf{0})$. If \mathbf{C} is nonsingular, then GCN possess a density function.

For a complete second-order description of the limiting distribution of any statistic $\hat{\boldsymbol{\theta}} \in \mathbb{C}^d$ we need to provide both the asymptotic covariance *and* the pseudo-covariance matrix. This may be clarified by noting that the real multivariate central limit theorem (e.g., [24, p. 385]) when written into a complex form reads as follows.

1) Complex Central Limit Theorem (CCLT): Let $\mathbf{z}_1, \dots, \mathbf{z}_n \in \mathbb{C}^k$ be i.i.d. random vectors from F with mean $\boldsymbol{\mu} \in \mathbb{C}^k$, finite covariance matrix $\mathbf{C}(F)$ and pseudo-covariance matrix $\mathcal{P}(F)$, then $\sqrt{n}(\bar{\mathbf{z}} - \boldsymbol{\mu}) \xrightarrow{L} \text{CN}_k(\mathbf{C}, \mathcal{P})$.

Estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta} \in \mathbb{C}^d$ based on i.i.d. random sample $\mathbf{z}_1, \dots, \mathbf{z}_n$ from F has *asymptotic GCN distribution* with asymptotic covariance matrix $\text{ASC}(\hat{\boldsymbol{\theta}}; F)$ and asymptotic pseudo-covariance matrix $\text{ASP}(\hat{\boldsymbol{\theta}}; F)$, if

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{L} \text{CN}_d(\text{ASC}(\hat{\boldsymbol{\theta}}; F), \text{ASP}(\hat{\boldsymbol{\theta}}; F)).$$

If $\text{ASP}(\hat{\boldsymbol{\theta}}; F) = \mathbf{0}$, then $\hat{\boldsymbol{\theta}}$ has *asymptotic CN distribution*. By CCLT, $\bar{\mathbf{z}}$ has GCN distribution with $\text{ASC}(\bar{\mathbf{z}}; F) = \mathbf{C}(F)$ and $\text{ASP}(\bar{\mathbf{z}}; F) = \mathcal{P}(F)$. Moreover, $\bar{\mathbf{z}}$ has asymptotic CN distribution if and only if F is second-order circular.

If a functional \mathbf{T} corresponding to an estimator $\hat{\boldsymbol{\theta}} = \mathbf{T}(F_n) \in \mathbb{C}^d$ is sufficiently regular and $\mathbf{z}_1, \dots, \mathbf{z}_n$ is an i.i.d. random sample from F , one has that [5], [18]

$$\sqrt{n}\{\hat{\boldsymbol{\theta}} - \mathbf{T}(F)\} = \sqrt{n} \left[\frac{1}{n} \sum_{i=1}^n \text{IF}\{\mathbf{z}_i; \mathbf{T}, F\} \right] + o_p(1). \quad (15)$$

It turns out that $E[\text{IF}(\mathbf{z}; \mathbf{T}, F)] = \mathbf{0}$ and, hence by CCLT, $\hat{\boldsymbol{\theta}}$ is asymptotically complex normal:

$$\sqrt{n}\{\hat{\boldsymbol{\theta}} - \mathbf{T}(F)\} \xrightarrow{L} \text{CN}_d(\text{ASC}(\hat{\boldsymbol{\theta}}; F), \text{ASP}(\hat{\boldsymbol{\theta}}; F)),$$

with

$$\text{ASC}(\hat{\boldsymbol{\theta}}; F) = E [\text{IF}(\mathbf{z}; \mathbf{T}, F)\text{IF}(\mathbf{z}; \mathbf{T}, F)^H] \quad (16)$$

$$\text{ASP}(\hat{\boldsymbol{\theta}}; F) = E [\text{IF}(\mathbf{z}; \mathbf{T}, F)\text{IF}(\mathbf{z}; \mathbf{T}, F)^T]. \quad (17)$$

Although (15) is often true, a rigorous proof may be difficult. However, given the form of the IF, (16) and (17) can be used to calculate an expression for the asymptotic covariance matrix and pseudo-covariance matrix of the estimator $\hat{\boldsymbol{\theta}}$ in a heuristic manner.

Adopting (16) and (17) as the definitions of ASC and ASP, it was shown in [25] that for the off-diagonal element \hat{c}_{ij} ($i \neq j$) of scatter matrix estimator $\hat{\mathbf{C}}$, one has that

$$\text{ASC}(\hat{c}_{ij}; F_{\Sigma}) = \tau_{\mathbf{C}} \sigma_{ii} \sigma_{jj} + \kappa_{\mathbf{C}} |\sigma_{ij}|^2$$

and

$$\text{ASP}(\hat{c}_{ij}; F_{\Sigma}) = (\tau_{\mathbf{C}} + \kappa_{\mathbf{C}}) \sigma_{ij}^2$$

where σ_{ij} denotes (i, j) th element of Σ , $\tau_{\mathbf{C}} = \tau_{\mathbf{C}}(t)$ and $\kappa_{\mathbf{C}} = \kappa_{\mathbf{C}}(t)$ are constants

$$\tau_{\mathbf{C}} = \frac{E[\alpha_{\mathbf{C}}^2(t)]}{k(k+1)} \quad \text{and} \quad \kappa_{\mathbf{C}} = E[\beta_{\mathbf{C}}^2(t)] - \frac{\tau_{\mathbf{C}}}{k} \quad (18)$$

which depend on the underlying CES distribution F_{Σ} via r.v.a. t possessing density (6). It will be shown in Section V that the asymptotic covariance structure of SMB-MVDR beamformer relies upon the constant $\tau_{\mathbf{C}}$, but not on $\kappa_{\mathbf{C}}$.

V. STATISTICAL ANALYSIS OF SMB-MVDR BEAMFORMERS

A. Influence Function

Based on Lemma 1, it is now possible to derive a general expression for the IF of the SMB-MVDR functional $\mathbf{w}_{\mathbf{C}}$.

Theorem 1: With the notation of Lemma 1, the influence function of SMB-MVDR functional $\mathbf{w}_{\mathbf{C}}$ at a CES distribution $F_{\Sigma} = \text{CE}_k(\Sigma, g)$ is given by

$$\text{IF}(\mathbf{t}; \mathbf{w}_{\mathbf{C}}, F_{\Sigma}) = \frac{\alpha_{\mathbf{C}}(r^2)}{\sigma_{\mathbf{C}}} [\mathbf{w}_{\mathbf{C}}^H - I](\mathbf{A}^{-1})^H \mathbf{u} \mathbf{u}^H \mathbf{A}^H \mathbf{w}_{\mathbf{C}}$$

where $\mathbf{w} = \mathbf{W}g$ is defined in (13) and $\sigma_{\mathbf{C}}$ is the scalar factor (7).

Remark 1: Theorem 1 shows that the IF of $\mathbf{w}_{\mathbf{C}}$ is continuous and bounded if $\alpha_{\mathbf{C}}(\cdot)$ is continuous and bounded. This follows by noting that when $\|\mathbf{t}\|$, or equivalently $r = \|\mathbf{A}^{-1}\mathbf{t}\|$, grows to infinity, $\mathbf{u} = \mathbf{A}^{-1}\mathbf{t}/r$ remains bounded. Hence, to validate qualitative robustness of SMB-MVDR beamformers employing scatter matrix \mathbf{C} we only need to study the $\alpha_{\mathbf{C}}(\cdot)$ function appearing in the IF of \mathbf{C} in Lemma 1. Note also that the IF of $\mathbf{w}_{\mathbf{C}}$ do not depend on the function $\beta_{\mathbf{C}}(\cdot)$, which together with $\alpha_{\mathbf{C}}(\cdot)$, fully determines the IF of \mathbf{C} . This is in line with result (8.2) of [26] obtained in the real case.

Remark 2: Theorem 1 also shows that $\text{IF}(\mathbf{g}; \mathbf{w}_{\mathbf{C}}, F_{\Sigma}) = \mathbf{0}$, i.e., if the contamination point \mathbf{t} equals the array response \mathbf{g} , then it causes zero influence on the functional.

For the covariance matrix \mathbf{C} , $\alpha_{\mathbf{C}}(x) = x$ and $\sigma_{\mathbf{C}} = E(t)/k$ with t having density (6). The IF of the associated SMB-MVDR functional $\mathbf{w}_{\mathbf{C}}$ is thus quadratic in $r = \|\mathbf{A}^{-1}\mathbf{t}\|$ and consequently unbounded. In [14], $\alpha_{\mathbf{C}}(\cdot)$ function of the M -functional of scatter \mathbf{C}_{φ} was found to be

$$\alpha_{\varphi}(x) = \frac{\varphi\left(\frac{x}{\sigma_{\varphi}}\right) x}{1 + c_{\varphi}} \quad (19)$$

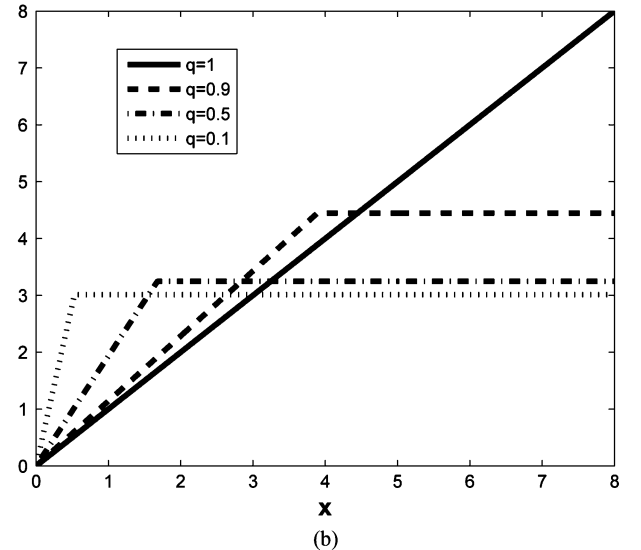
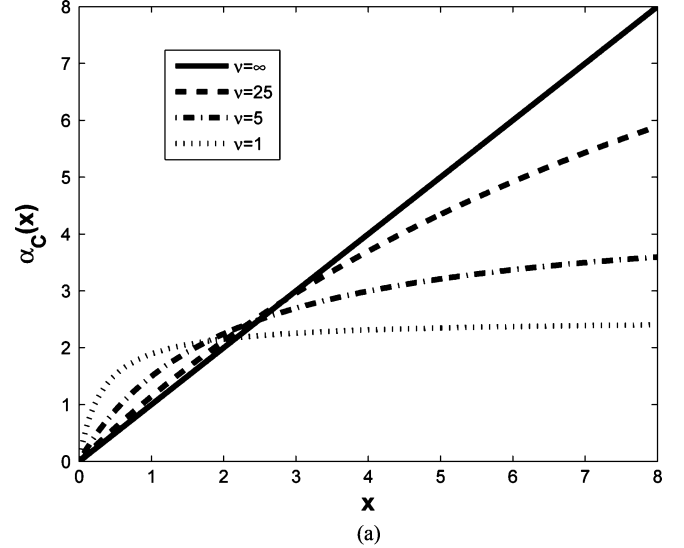


Fig. 2. $\alpha_{\varphi}(x)$ functions of (a) MLT(ν) estimators and (b) HUB(q) estimators at Φ_2 .

where the constant c_{φ} is defined as

$$c_{\varphi} = \frac{1}{k(k+1)} E\left[\varphi'\left(\frac{t}{\sigma_{\varphi}}\right) \frac{t^2}{\sigma_{\varphi}^2}\right] \quad (20)$$

where t has p.d.f. (6) and the scalar factor σ_{φ} is the solution to (12). Next corollary then follows at once from (19) and Theorem 1.

Corollary 1: The influence function of SMB-MVDR weight \mathbf{w}_{φ} based on M -functional of scatter \mathbf{C}_{φ} is continuous and bounded if and only if $\varphi(x)x$ is continuous and bounded.

Fig. 2(a) depicts $\alpha_{\varphi}(x)$ of MLT(ν) functionals for choices $\nu = 1, 5, 25, \infty$ at the Φ_2 distribution. As we can see, α_{φ} -functions corresponding to the robust MLT(1) and MLT(5) functionals are bounded and continuous. However, when ν increases, α_{φ} -function resembles more and more straight line corresponding to α_{φ} function of the covariance matrix (i.e., $\nu = \infty$). Fig. 2(b) depicts $\alpha_{\varphi}(x)$ of HUB(q) functionals for choices $q = 1, 0.9, 0.5, 0.1$ at the Φ_2 distribution. Recall that HUB(1) correspond to the covariance matrix. We observe that

TABLE I
ASYMPTOTIC RELATIVE EFFICIENCIES OF $\hat{\mathbf{w}}_\varphi$ BASED ON HUB(q) AND MLT(ν) ESTIMATORS AT Φ_k , $T_{k,5}$ AND $T_{k,1}$ DISTRIBUTIONS (AND DIMENSIONS $k = 2, 3, 5, 10$). RECALL THAT HUB (1) CORRESPOND TO SCM

$k =$	Normal (Φ_k)				$T_{k,5}$				Cauchy ($T_{k,1}$)			
	2	3	5	10	2	3	5	10	2	3	5	10
HUB(1)	1	1	1	1	.407	.394	.378	.360	0	0	0	0
HUB(.9)	.961	.975	.986	.994	.963	.952	.940	.933	.770	.779	.798	.833
HUB(.5)	.823	.876	.923	.961	.949	.974	.986	.988	.975	.969	.964	.964
HUB(.1)	.698	.777	.853	.921	.849	.912	.960	.988	.963	.983	.993	.997
MLT(5)	.906	.910	.921	.941	1	1	1	1	.827	.877	.927	.969
MLT(2)	.825	.849	.882	.925	.968	.977	.987	.995	.976	.985	.992	.997
MLT(1)	.769	.811	.861	.917	.923	.948	.972	.990	1	1	1	1

the robust HUB(q) functionals with $q \leq 0.9$ possess bounded and continuous α_φ -functions.

B. Asymptotic Relative Efficiencies

If we take (16) and (17) as the definitions of the asymptotic covariance matrix and pseudo-covariance matrix of a functional then the next theorem holds.

Theorem 2: The asymptotic covariance matrix of the estimated SMB-MVDR weight $\hat{\mathbf{w}}_C$ when sampling from $F_\Sigma = CE_k(\Sigma, g)$ is

$$ASC(\hat{\mathbf{w}}_C; F_\Sigma) = \lambda_C (\mathbf{W} - \mathbf{w}\mathbf{w}^H)$$

where $\mathbf{w} = \mathbf{W}\mathbf{g}$ is defined in (13) and $\lambda_C = \lambda_C(t)$ is a positive constant scalar, $\lambda_C = \tau_C/\sigma_C^2$, where τ_C is given in (18) and σ_C is the scalar factor (7). Furthermore, the asymptotic pseudo-covariance matrix of $\hat{\mathbf{w}}_C$ vanishes, i.e., $ASP(\hat{\mathbf{w}}_C; F_\Sigma) = \mathbf{0}$.

Remark 3: Note that the ASC of $\hat{\mathbf{w}}_C$ depends on the selected scatter matrix \mathbf{C} and on the functional form of the CES distribution F_Σ only via the real-valued positive multiplicative constant λ_C . (Observe that the matrix term $\mathbf{W} - \mathbf{w}\mathbf{w}^H$ do not depend on the choice of \mathbf{C} and on F_Σ only via Σ .) Hence, comparisons of this single scalar index is needed only. It is a suprising result that ASP vanishes, which means that $\hat{\mathbf{w}}_C$ has asymptotic CN distribution.

Remark 4: Note that $ASC(\hat{\mathbf{w}}_C; F_\Sigma)$ is singular and of rank $k - 1$ (since the nullspace of $\mathbf{W} - \mathbf{w}\mathbf{w}^H$ has dimension 1 due to MVDR constraint $\mathbf{w}^H\mathbf{g} = 1$, so $(\mathbf{W} - \mathbf{w}\mathbf{w}^H)\mathbf{g} = \mathbf{0}$). Thus, the asymptotic CN distribution of $\hat{\mathbf{w}}_C$ is singular. This is expected result since singular distributions commonly arise in constrained parameter estimation problems, where the constraint imposes certain degree of deterministic to the estimator.

The ARE of $\hat{\mathbf{w}}_C$ is calculated as

$$ARE(\hat{\mathbf{w}}_C; F_\Sigma) = \frac{\text{Tr}[ASC(\hat{\mathbf{w}}_{\text{mle}}; F_\Sigma)]}{\text{Tr}[ASC(\hat{\mathbf{w}}_C; F_\Sigma)]} = \frac{\lambda_{\text{mle}}}{\lambda_C}$$

where λ_{mle} correspond to λ_C value associated with the optimal SMB-MVDR beamformer $\hat{\mathbf{w}}_{\text{mle}}$. Thus, $ARE(\hat{\mathbf{w}}_C; F_\Sigma) \leq 1$. Loosely speaking, if $\hat{\mathbf{w}}_C$ is based on a sample of n observations (and n is large), then $n \cdot ARE(\hat{\mathbf{w}}_C; F_\Sigma)$ is the sample size needed for $\hat{\mathbf{w}}_{\text{mle}}$ to achieve the same accuracy as $\hat{\mathbf{w}}_C$.

For $\hat{\mathbf{w}}_\varphi$, the constant λ_C , denoted by λ_φ , is

$$\lambda_\varphi(t) = \frac{E[\varphi^2 \left(\frac{t}{\sigma_\varphi}\right) \left(\frac{t}{\sigma_\varphi}\right)^2]}{k(k+1)(1+c_\varphi)^2}$$

where the quadratic form $t = \mathbf{z}^H \Sigma^{-1} \mathbf{z}$ possess density (6), c_φ is given in (20) and σ_φ is the solution to (12). For $\hat{\mathbf{w}}_C$ (i.e., $\varphi \equiv 1$), we have that

$$\lambda_C(t) = \frac{E(t^2)}{k(k+1)} = \begin{cases} \lambda_{\text{mle}}(t) = 1, & \text{under } \Phi_k \\ \frac{(\nu-2)}{(\nu-4)}, & \text{under } T_{k,\nu}. \end{cases}$$

Since $\lambda_C(t)$ depends on the second-order moment $E(t^2)$, the underlying CES distribution F_Σ needs to possess finite fourth-order moments in order that $\hat{\mathbf{w}}_C$ possess asymptotic CN distribution. For example, at $T_{k,\nu}$ distribution with $\nu < 5$, $\hat{\mathbf{w}}_C$ do not possess limiting CN distribution.

Table I reports the AREs of $\hat{\mathbf{w}}_\varphi$ based on HUB(q) (using $q = 1, 0.9, 0.5, 0.1$) and MLT(ν) (using $\nu = 1, 2, 5$) under complex normal (Φ_k), complex Cauchy ($T_{k,1}$) and $T_{k,5}$ distributions for some choices of dimension k . Recall that HUB(1) correspond to the SCM. At CN distribution, we note that the HUB(0.9) has the best performance among its robust alternatives and that the efficiencies associated with HUB(q) and MLT(ν) are increasing with the dimension: for example, at $k = 10$, HUB(0.9) experiences only 0.006% efficiency loss and MLT(1) which ranks the lowest, has a moderate 8.3% loss in efficiency. Hence, adding more sensors to the array increases the (asymptotic) efficiency of the estimated SMB-MVDR beamformers based upon the above robust M -estimators. At $T_{k,5}$ distribution, employed M -estimators are superior to the conventional covariance matrix based beamformer. Curiously, the efficiencies for the covariance matrix and HUB(0.9) are decreasing with the dimension (the decrease being faster for the covariance matrix) whereas the efficiencies for HUB(0.5), HUB(0.1), MLT(1) and MLT(2) are increasing and tending towards the optimal value 1 as the dimensionality grows. At complex Cauchy distribution, all the robust M -estimators are performing very well and their efficiencies are increasing with the dimension. To conclude, these asymptotic efficiencies clearly favor estimated SMB-MVDR beamformers based upon robust M -estimators of scatter, since they combine a high efficiency with appealing robustness properties.

Finite sample covariance matrix (among many other results) of the conventional estimated MVDR beamformer weight $\hat{\mathbf{w}}_C$ was found in [27, p. 1785] at Φ_k to be

$$\mathbf{C}(\hat{\mathbf{w}}_C) = \frac{1}{n-k+1} (\mathbf{W} - \mathbf{w}\mathbf{w}^H).$$

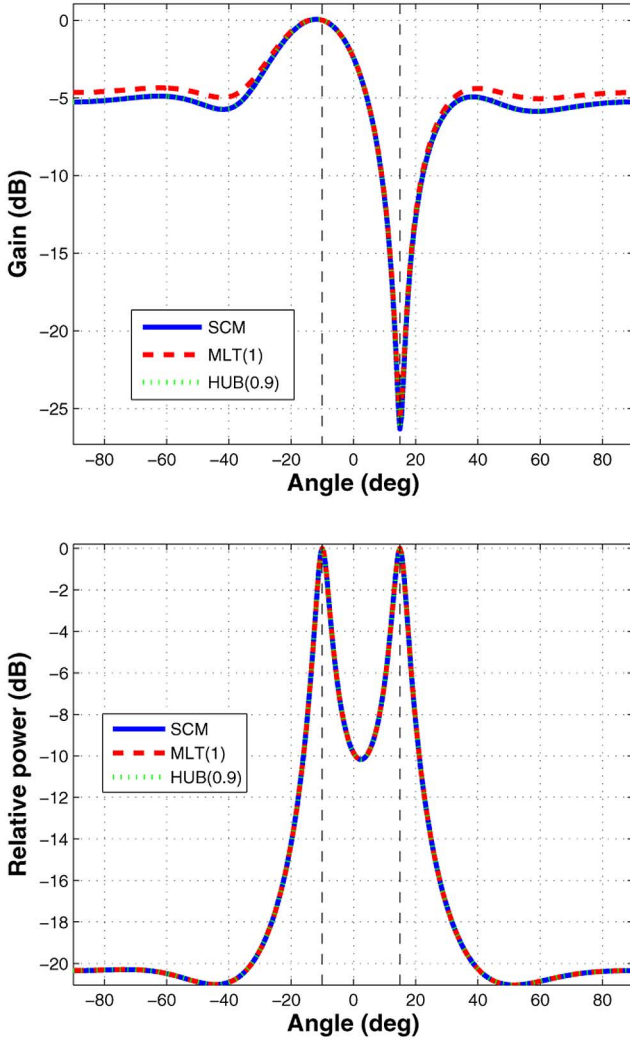


Fig. 3. Averaged SMB-MVDR beampatterns (left) and spectrums (right) for Setting A ($n = 500$, SOI at -10° , interferer at 15°). All estimators perform comparably.

Hence, we observe the expected result: $\mathcal{C}(\sqrt{n}\hat{\mathbf{w}}_{\mathcal{C}}) \rightarrow \text{ASC}(\hat{\mathbf{w}}_{\mathcal{C}}; \Phi_k) = (\mathbf{W} - \mathbf{w}\mathbf{w}^H)$ as $n \rightarrow \infty$. The results of [27] were shown to apply for a data matrix $Z = (\mathbf{z}_1 \ \mathbf{z}_2 \ \cdots \ \mathbf{z}_n) \in \mathbb{C}^{k \times n}$ possessing a matrix-variate CES distribution. A disadvantage of matrix-variate CES distribution is that it does not allow $\mathbf{z}_1, \dots, \mathbf{z}_n$ to be i.i.d. from F_{Σ} other than Φ_k . Since i.i.d.'ness is a key assumption in this paper, the class of matrix-variate CES distributions is not a permissible model here.

VI. NUMERICAL EXAMPLES

A. Finite Sample Robustness

As noted earlier, influence function is a theoretical (asymptotic) tool to quantify the robustness of the functional form of an estimator. We now consider *finite sample* robustness of the SMB-MVDR beamformers on simulated data sets. The simulation settings are described below.

Setting A: Gaussian noise: A four-sensor uniform linear array ($\lambda/2$ spacing) receives two uncorrelated circular normal

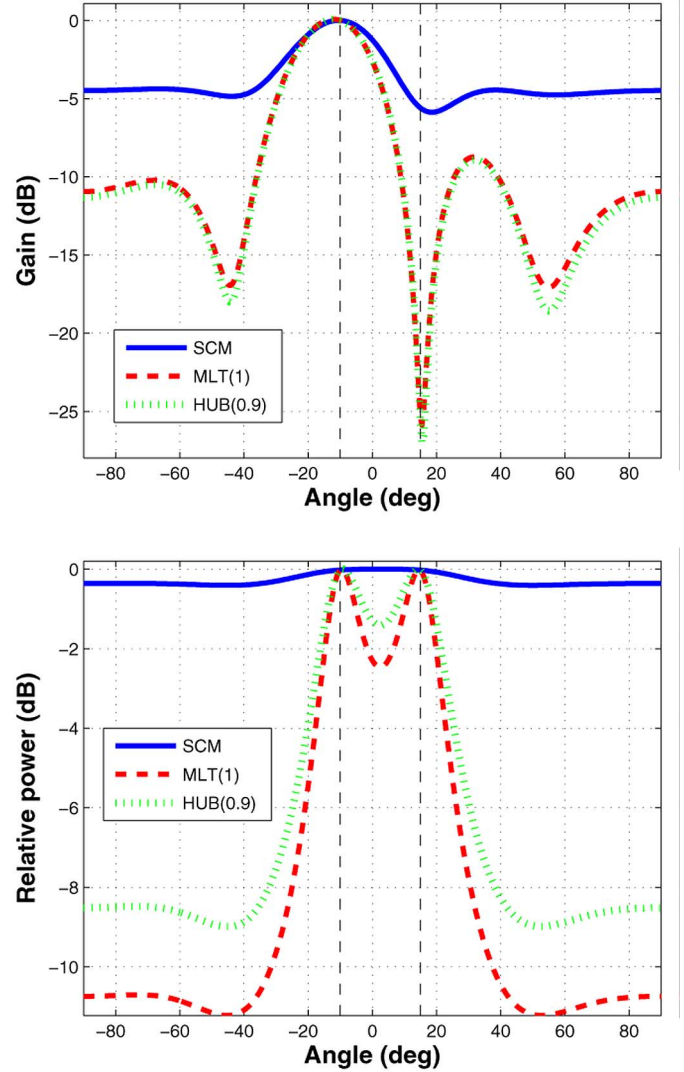


Fig. 4. Averaged SMB-MVDR beampatterns (top) and spectrums (bottom) for Setting B ($n = 500$, SOI at -10° , interferer at 15°). SCM fails, but robust HUB(0.9) and MLT(1) estimators perform very well.

(Gaussian) signals (of equal variance σ_s^2) with direction-of-arrivals (DOAs) at -10° (signal-of-interest, SOI) and 15° (interferer). The array output is corrupted by additive noise vector (independent of the signals), whose elements are i.i.d. following (circular) complex Gaussian distribution with scale parameter σ_n^2 .

Setting B: Cauchy noise: As Setting A, but noise vector has i.i.d. elements following circular complex Cauchy distribution with scale σ_n^2 . In both settings, the SNR (dB) is $10 \log_{10}[\sigma_s^2/\sigma_n^2] = 15$ dB.

Figs. 3 and 4 depict the estimated SMB-MVDR beampatterns and spectrums (number of snapshots $n = 500$) for look direction -10° for Settings A and B averaged over 100 realizations. The employed scatter matrices are the SCM [i.e., HUB(1)], MLT(1) and HUB(0.9). In the Gaussian noise case (Setting A; Fig. 3), the estimated beampatterns are closely similar, in fact, overlapping for the SCM and HUB(0.9). The estimated spectrums associated with the different estimators are overlapping in the Gaussian case, so they provide essentially the same DOA

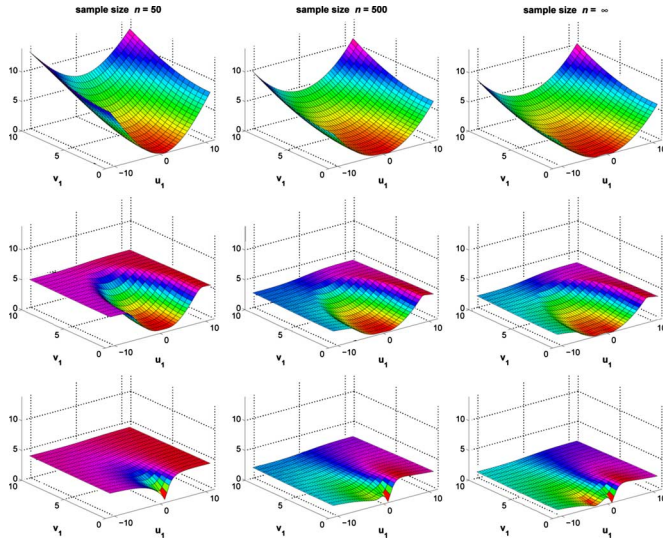


Fig. 5. Averaged $\|\text{EIF}(\mathbf{t}; \hat{\mathbf{w}}_{\mathbf{C}}, Z_n)\|$ for the SCM (top row), HUB(0.9) (middle row) and MLT(1) (bottom row). Samples Z_n ($n = 50, 500$) are generated from model of Setting A; $n = \infty$ correspond to the plot of $\|\text{IF}(\mathbf{t}; \mathbf{w}_{\mathbf{C}}, F_{\Sigma})\|$. The contamination point $\mathbf{t} = (t_1, t_2)^T$ is such that t_2 is fixed and $t_1 = u_1 + jv_1 \in \mathbb{C}$ varies.

estimates. In the Cauchy noise case (Setting B; Fig. 4), however, the conventional MVDR fails completely and can not resolve the two sources: the estimated beampattern and the spectrum are flat and the mainlobe and the peaks cannot be well identified. The beampattern associated with MLT(1) and HUB(0.9), however, has a narrow mainlobe centered at the look direction and places a deep null in the direction of interference. Also the spectrums for MLT(1) and HUB(0.9) shows two sharp peaks at the DOAs of the sources. Hence, the performance loss is negligible by employing MLT(1) or HUB(0.9) instead of the SCM in nominal Gaussian noise conditions. However, significant gain in performance is obtained when the noise is heavy-tailed Cauchy.

We now compute the empirical influence functions (EIFs) for the above estimators for simulated data sets Z_n from Setting A. In this case, $\mathbf{z} \sim \mathcal{CN}_4(\Sigma)$ with $\Sigma = \sigma_s^2 \mathbf{G}\mathbf{G}^H + \sigma_n^2 \mathbf{I}$, where $\mathbf{G} = (\mathbf{g} \quad \mathbf{g}_1) \in \mathbb{C}^{4 \times 2}$ denotes the array response matrix of ULA for DOAs at -10° (SOI) and 15° (interferer). Let the $k = 4$ -variate contaminating vector \mathbf{t} be such that only the first component $t_1 = u_1 + jv_1$ is allowed to vary, and the remaining components have fixed values: $t_i = g_i$, $i \in \{2, 3, 4\}$, where g_i denotes the i th component of the array response \mathbf{g} . An informative picture on the effect of contamination $t_1 = x_1 + jy_1$ on $\hat{\mathbf{w}}_{\mathbf{C}}$ is obtained by the surface plot of the norm of the empirical influence function $\|\text{EIF}(\mathbf{t}; \hat{\mathbf{w}}_{\mathbf{C}}, Z_n)\|$ with respect to u_1 and v_1 . The EIFs in Fig. 5 are averages over 100 realizations. Sample lengths are $n = 50, 500, \infty$, where the surface plots under $n = \infty$ correspond to asymptotic value $\|\text{IF}(\mathbf{t}; \mathbf{w}_{\mathbf{C}}, \Phi_4)\|$. As expected, we observe that when the sample size grows (from $n = 50$ to $n = 500$), the calculated EIF surfaces more accurately resemble the corresponding theoretical IF surface. However, at the small sample size ($n = 50$), the relative influence of an additional observation on the estimator is a bit larger than what the IF would indicate. The surface plots neatly demonstrate the non-robustness of the conventional MVDR beamformer for both the finite and large sample cases: outlying points with large values of

x_1 and/or y_1 have bounded influence in case of HUB(0.9) or MLT(1) but large and unbounded influence when the conventional SCM is employed.

B. Finite Sample Efficiencies

The asymptotic behavior is often used to approximate the small sample behavior of an estimator. Therefore, it is important to investigate the accuracies of these approximations. In the following simulation study, estimated finite sample relative efficiencies are compared with the asymptotic numbers presented in Table I.

We generated $M = 1000$ samples ($n = 50, 100, 300$) from k -variate ($k = 5, 10$) complex normal (Φ_k), Cauchy ($T_{k,1}$) and $T_{k,5}$ distributions with

$$\Sigma = \sigma^2 \mathbf{g}\mathbf{g}^H + \Omega \quad (21)$$

where $\Omega \in \text{PDH}(k)$ with $\text{Tr}(\Omega) = k$ and $\sigma^2 > 0$. Since Σ is proportional to covariance matrix (provided that it exists), (21) represents the decomposition of Σ (and thus of \mathbf{C}) into signal and noise-plus-interference components. Furthermore, since the noise-plus-interference matrix term Ω is normalized, σ^2 represents the signal-to-interference-plus-noise ratio (SINR) averaged across the k sensors. In our simulation, $\text{SINR} = 10 \log_{10} \sigma^2 = 20$ (dB), Ω was generated randomly and \mathbf{g} is the array response of the ULA at 10° (of SOI). Obtained simulation results were similar for any other choices of σ^2 , Ω or \mathbf{g} . The mean-squared error (MSE) of the SMB-MVDR weight $\hat{\mathbf{w}}_{\mathbf{C}}$ was calculated by $\text{MSE}(\hat{\mathbf{w}}_{\mathbf{C}}) = \text{Tr}\{1/M \sum_{m=1}^M (\hat{\mathbf{w}}_{\mathbf{C}}^{(m)} - \mathbf{w})(\hat{\mathbf{w}}_{\mathbf{C}}^{(m)} - \mathbf{w})^H\}$, where $\hat{\mathbf{w}}_{\mathbf{C}}^{(m)}$ denotes the computed value for the m th generated sample. Finite sample relative efficiency of $\hat{\mathbf{w}}_{\mathbf{C}}$ listed in Table II is calculated as the ratio $\text{MSE}(\hat{\mathbf{w}}_{\text{mle}})/\text{MSE}(\hat{\mathbf{w}}_{\mathbf{C}})$. The corresponding AREs from Table I are also listed under $n = \infty$ for easy reference. Recall that HUB(1) correspond to the SCM. The results show that (except for the SCM at $T_{k,5}$) the finite sample relative efficiencies approximate very well the AREs even at small sample sizes. At $T_{k,5}$, the SCM appears to have a gain in efficiency at small sample sizes since the estimated finite sample efficiencies are considerably higher than the ARE; also the convergence of finite sample efficiencies toward the ARE as n increases is a bit slow.

VII. CONCLUSION

Optimal array processors are often scale-invariant functions of the covariance matrix \mathbf{C} . Hence, they can be based on generalized covariance matrix \mathbf{C} , called the scatter matrix. Scatter matrix is a more general notion than covariance matrix, but proportional to it under the class of CES distributions. In addition to MVDR considered in this paper, also many other array processors, such as the classical complex least-squares estimator [12, p. 522, 529–530] satisfy (1). Also many high-resolution subspace methods such as MUSIC and ESPRIT (see [17]), are scale-invariant since they depend on the covariance matrix only through its eigenvectors. In this paper, we focused on a scatter matrix based (SMB-)MVDR beamformer weight vector and derived a general expression for its influence function and asymptotic covariance structure.

TABLE II
FINITE SAMPLE RELATIVE EFFICIENCIES OF \hat{w}_φ BASED ON HUB(q) AND MLT(ν) ESTIMATORS. SAMPLES (LENGTHS $n = 50, 300, 1000$) ARE SIMULATED FROM Φ_k AND $T_{k,5}$ DISTRIBUTIONS (DIMENSIONS $k = 5, 10$)

$n =$	Normal (Φ_k)				$T_{k,5}$				Cauchy ($T_{k,1}$)			
	50	100	300	∞	50	100	300	∞	50	100	300	∞
$k = 5$												
HUB(1)	1	1	1	1	.666	.573	.497	.378	.090	.047	.016	0
HUB(.9)	.992	.989	.988	.986	.936	.934	.936	.940	.782	.788	.782	.798
HUB(.5)	.934	.929	.927	.923	.979	.985	.985	.986	.956	.955	.952	.964
HUB(.1)	.863	.857	.860	.853	.960	.964	.957	.960	.996	.992	.995	.993
MLT(5)	.934	.926	.926	.921	1	1	1	1	.915	.921	.917	.927
MLT(1)	.873	.867	.869	.861	.971	.976	.970	.972	1	1	1	1
$k = 10$												
HUB(1)	1	1	1	1	.736	.638	.531	.360	0.185	.094	.031	0
HUB(.9)	.994	.991	.995	.994	.935	.932	.921	.933	.812	.819	.823	.833
HUB(.5)	.959	.950	.961	.961	.990	.987	.985	.988	.964	.958	.967	.964
HUB(.1)	.916	.905	.919	.921	.985	.986	.992	.988	.998	.998	.996	.997
MLT(5)	.943	.929	.939	.941	1	1	1	1	.965	.962	.967	.969
MLT(1)	.916	.902	.915	.917	.988	.989	.994	.990	1	1	1	1

We showed that the conventional estimated MVDR beamformer (employing the SCM) is not robust and has a serious loss in efficiency if the traditional assumption of CN distribution is not valid. On the other hand, SMB-MVDR beamformers employing the considered robust M -estimators of scatter were shown to be statistically robust (possessing a bounded and continuous IF) and have good efficiencies under CN distribution and (heavy-tailed) complex $T_{k,\nu}$ distribution, e.g., at complex Cauchy distribution ($T_{k,1}$). For example, the Huber's M -estimator HUB(q) with $q = 0.9$ is a very safe choice, as it suffers negligible performance loss compared to SCM under traditional CN assumption, but has a superior performance for heavy-tailed distributions. On the other hand, MLT(1) gives the best safeguard against outliers and has a moderate efficiency loss under CN distribution. Our examples with simulated data confirmed the theoretical findings. To conclude, our findings favor beamformers based upon M -estimators of scatter, since they combine a high efficiency with appealing statistical robustness properties.

APPENDIX A

COMPUTATION OF M -ESTIMATOR OF SCATTER

Given any initial estimate $\hat{\mathbf{C}}_0 \in \text{PDH}(k)$, the iterations

$$\hat{\mathbf{C}}_{m+1} = \frac{1}{n} \sum_{i=1}^n \varphi(\mathbf{z}_i^H \hat{\mathbf{C}}_m^{-1} \mathbf{z}_i) \mathbf{z}_i \mathbf{z}_i^H$$

converge to the solution $\hat{\mathbf{C}}_\varphi$ of (10) under some mild regularity conditions. The authors of [2], [18], and [28] consider the real case only, but the complex case follows similarly. See also discussions in [3].

As an example, let the initial estimate be the SCM, i.e., $\hat{\mathbf{C}}_0 = \hat{\mathbf{C}}$. The first iteration, or the "1-step M -estimator," is simply a weighed sample covariance matrix

$$\hat{\mathbf{C}}_1 = \frac{1}{n} \sum_{i=1}^n w_i \mathbf{z}_i \mathbf{z}_i^H, \quad w_i = \varphi(\mathbf{z}_i^H \hat{\mathbf{C}}^{-1} \mathbf{z}_i).$$

If $\varphi(\cdot)$ is a robust weighting function, then $\hat{\mathbf{C}}_1$ is a robustified version of $\hat{\mathbf{C}}$. At the second iteration step, we calculate $\hat{\mathbf{C}}_2$ as a weighted sample covariance matrix using weights $w_i = \varphi(\mathbf{z}_i^H \hat{\mathbf{C}}_1^{-1} \mathbf{z}_i)$ and proceed analogously until the iterations $\hat{\mathbf{C}}_3, \hat{\mathbf{C}}_4, \dots$ "converge," i.e., $\|\mathbf{I} - \hat{\mathbf{C}}_{m-1}^{-1} \hat{\mathbf{C}}_m\| < \varepsilon$, where $\|\cdot\|$ is a matrix norm and ε is predetermined tolerance level, e.g., $\varepsilon = 0.001$. To reduce computation time, one can always stop after m (e.g., $m = 4$) iterations and take the "m-step M -estimator" $\hat{\mathbf{C}}_m$ as an approximation for the true M -estimator $\hat{\mathbf{C}}_\varphi$. MATLAB functions to compute MLT(ν) and HUB(q) estimators are available at <http://wooster.hut.fi/~esollila/MVDR/>.

APPENDIX B

PROOF OF THEOREM 1

Lemma 2: With the notation of Lemma 1, the influence function of \mathbf{C}^{-1} at $F_\Sigma = \text{CE}_k(\Sigma, g)$ is given by

$$\text{IF}(\mathbf{t}; \mathbf{C}^{-1}, F_\Sigma) = -\sigma_{\mathbf{C}}^{-2} \{ \alpha_{\mathbf{C}}(r^2) \mathbf{A}^{-H} (\mathbf{u} \mathbf{u}^H - k^{-1} \mathbf{I}) \mathbf{A}^{-1} + \beta_{\mathbf{C}}(r^2) \Sigma^{-1} \}.$$

Proof: Write $F_{\varepsilon, \mathbf{t}} = (1 - \varepsilon) F_\Sigma + \varepsilon \Delta_{\mathbf{t}}$ for the ε contaminated distribution at \mathbf{t} . Note that $F_{0, \mathbf{t}} = F_\Sigma$. Since $\mathbf{C}(F_{\varepsilon, \mathbf{t}}) \mathbf{C}^{-1}(F_{\varepsilon, \mathbf{t}}) = \mathbf{I}$, we obtain that

$$\mathbf{0} = \frac{\partial}{\partial \varepsilon} \left\{ \mathbf{C}(F_{\varepsilon, \mathbf{t}}) \mathbf{C}^{-1}(F_{\varepsilon, \mathbf{t}}) - \mathbf{I} \right\} \Big|_{\varepsilon=0} = \text{IF}(\mathbf{t}; \mathbf{C}, F_\Sigma) \mathbf{C}^{-1}(F_\Sigma) + \mathbf{C}(F_\Sigma) \text{IF}(\mathbf{t}; \mathbf{C}^{-1}, F_\Sigma).$$

Rearranging the terms and recalling that $\mathbf{C}(F_\Sigma) = \sigma_{\mathbf{C}} \Sigma$ gives

$$\text{IF}(\mathbf{t}; \mathbf{C}^{-1}, F_\Sigma) = -\sigma_{\mathbf{C}}^{-2} \Sigma^{-1} \text{IF}(\mathbf{t}; \mathbf{C}, F_\Sigma) \Sigma^{-1}$$

which, after substituting the expression for $\text{IF}(\mathbf{t}; \mathbf{C}, F_\Sigma)$ from Lemma 1, produces the stated expression for the influence function of \mathbf{C}^{-1} . \blacksquare

Proof of Theorem 1: Simple derivation yields

$$\begin{aligned} \text{IF}(\mathbf{t}; \mathbf{w}_C, F_\Sigma) &= \frac{\partial}{\partial \varepsilon} \mathbf{w}_C(F_{\varepsilon, \mathbf{t}}) \Big|_{\varepsilon=0} = \frac{\partial}{\partial \varepsilon} \frac{\mathbf{C}^{-1}(F_{\varepsilon, \mathbf{t}}) \mathbf{g}}{\mathbf{g}^H \mathbf{C}^{-1}(F_{\varepsilon, \mathbf{t}}) \mathbf{g}} \Big|_{\varepsilon=0} \\ &= \left(\frac{\partial}{\partial \varepsilon} \mathbf{C}^{-1}(F_{\varepsilon, \mathbf{t}}) \Big|_{\varepsilon=0} \right) \frac{\mathbf{g}}{\mathbf{g}^H \mathbf{C}^{-1}(F_\Sigma) \mathbf{g}} \\ &\quad + \mathbf{C}^{-1}(F_\Sigma) \mathbf{g} \frac{\partial}{\partial \varepsilon} [\mathbf{g}^H \mathbf{C}^{-1}(F_{\varepsilon, \mathbf{t}}) \mathbf{g}]^{-1} \Big|_{\varepsilon=0}. \end{aligned}$$

Then by noting that

$$\begin{aligned} \frac{\partial}{\partial \varepsilon} [\mathbf{g}^H \mathbf{C}^{-1}(F_{\varepsilon, \mathbf{t}}) \mathbf{g}]^{-1} \Big|_{\varepsilon=0} &= \\ &= -[\mathbf{g}^H \mathbf{C}^{-1}(F_\Sigma) \mathbf{g}]^{-2} \mathbf{g}^H \left(\frac{\partial}{\partial \varepsilon} \mathbf{C}^{-1}(F_{\varepsilon, \mathbf{t}}) \Big|_{\varepsilon=0} \right) \mathbf{g} \end{aligned}$$

and that $\mathbf{C}(F_\Sigma) = \sigma_C \Sigma$, the earlier equation can be simplified to a form

$$\text{IF}(\mathbf{t}; \mathbf{w}_C, F_\Sigma) = \frac{1}{\sigma_C} \left[\frac{\Sigma^{-1} \mathbf{g}}{\mathbf{g}^H \Sigma^{-1} \mathbf{g}} \mathbf{g}^H - \mathbf{I} \right] \left[-\sigma_C^2 \frac{\text{IF}(\mathbf{t}; \mathbf{C}^{-1}, F_\Sigma) \mathbf{g}}{\mathbf{g}^H \Sigma^{-1} \mathbf{g}} \right]. \quad (22)$$

Using Lemma 2 and noting that $\Sigma^{-1} = \mathbf{A}^{-H} \mathbf{A}^{-1}$ yields

$$\begin{aligned} &= -\sigma_C^2 \frac{\text{IF}(\mathbf{t}; \mathbf{C}^{-1}, F_\Sigma) \mathbf{g}}{\mathbf{g}^H \Sigma^{-1} \mathbf{g}} \\ &= \alpha_C(r^2) \mathbf{A}^{-H} (\mathbf{u} \mathbf{u}^H - k^{-1} \mathbf{I}) \mathbf{A}^H \mathbf{w} + \beta_C(r^2) \mathbf{w} \\ &= \alpha_C(r^2) \mathbf{A}^{-H} \mathbf{u} \mathbf{u}^H \mathbf{A}^H \mathbf{w} + \left[\beta_C(r^2) - \frac{\alpha_C(r^2)}{k} \right] \mathbf{w}. \end{aligned}$$

Substituting the previous equation into (22) and observing that $[\mathbf{w} \mathbf{g}^H - \mathbf{I}] \mathbf{w} = \mathbf{0}$ (as $\mathbf{g}^H \mathbf{w} = 1$), completes the proof. \square

APPENDIX C PROOF OF THEOREM 2

First, we need to introduce some notations. Let $\text{vec}(\cdot)$ denote an operator that transforms a matrix into a vector by stacking the columns of the matrix and \otimes denote the *Kronecker product*: for any matrix \mathbf{A} and \mathbf{B} , $\mathbf{A} \otimes \mathbf{B}$ is a block matrix with (i, j) -block being equal to $a_{ij} \mathbf{B}$. An important identity involving the Kronecker product and vec -operator is

$$\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B}) \quad (23)$$

where \mathbf{A} , \mathbf{B} and \mathbf{C} are matrices such that the product \mathbf{ABC} is properly defined. Also, some useful, well-known properties of Kronecker product are: $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$ (where matrices \mathbf{A} , \mathbf{B} , \mathbf{C} and \mathbf{D} are such that the products \mathbf{AC} and \mathbf{BD} are properly defined), $(\mathbf{A} \otimes \mathbf{B})^H = \mathbf{A}^H \otimes \mathbf{B}^H$, $a \otimes \mathbf{A} = a \mathbf{A}$ and $\mathbf{a}^T \otimes \mathbf{b} = \mathbf{b} \mathbf{a}^T$. A *commutation matrix* [29] $\mathbf{I}_{k,k}$ is a $k^2 \times k^2$ block matrix with (i, j) -block equal to a $k \times k$ matrix that has a 1 at entry (j, i) and 0's elsewhere. The matrix derives its name from the property

$$\mathbf{I}_{k,k} (\mathbf{A} \otimes \mathbf{B}) = (\mathbf{B} \otimes \mathbf{A}) \mathbf{I}_{k,k}. \quad (24)$$

Finally, we denote $\mathbb{C}S^k = \{\mathbf{z} \in \mathbb{C}^k : \|\mathbf{z}\| = 1\}$ the unit complex k -sphere.

In the proof of Theorem 2, we rely upon Lemmas 3 and 4.

Lemma 3 (Krishnaiah and Lin [1]): Let $\mathbf{z} \sim \text{CE}_k(\Sigma, g)$ with $\Sigma = \mathbf{A} \mathbf{A}^H$. Then $r = \|\mathbf{A}^{-1} \mathbf{z}\|$ and $\mathbf{u} = \mathbf{A}^{-1} \mathbf{z}/r$ are independent, \mathbf{u} has a uniform distribution on $\mathbb{C}S^k$.

Lemma 4: Suppose \mathbf{u} has a uniform distribution on $\mathbb{C}S^k$.

[a] The only fourth-order moments $E[\tilde{u}_i \tilde{u}_j \tilde{u}_k \tilde{u}_l]$ (where i, j, k, l are arbitrary integers between $1, \dots, k$ and \tilde{u}_i is by convention either u_i^* or u_i) not equal to zero are

$$E[|u_i|^4] = \frac{2}{k(k+1)} \quad \text{and} \quad E[|u_i|^2 |u_j|^2] = \frac{1}{k(k+1)}$$

for all $i, j \in \{1, \dots, k\}$, $i \neq j$.

[b] Write $\mathbf{v} = \text{vec}(\mathbf{u} \mathbf{u}^H)$. Then $E[\mathbf{v} \mathbf{v}^H] = [k(k+1)]^{-1} (\mathbf{I} + \text{vec}(\mathbf{I}) \text{vec}(\mathbf{I})^T)$ and $E[\mathbf{v} \mathbf{v}^T] = [k(k+1)]^{-1} (\mathbf{I}_{k,k} + \text{vec}(\mathbf{I}) \text{vec}(\mathbf{I})^T)$.

Proof: [a] The distribution of \mathbf{u} remains invariant under unitary linear transformations, i.e., $\mathbf{u} \stackrel{d}{=} \mathbf{P} \mathbf{u}$ for all unitary matrices \mathbf{P} (so $\mathbf{P}^H \mathbf{P} = \mathbf{I}$), e.g., we may choose \mathbf{P} as $\mathbf{P} = \text{diag}(\exp(j\theta_1), \dots, \exp(j\theta_k))$ for arbitrary $\theta_i \in \mathbb{R}$, $i = 1, \dots, k$. This implies, e.g., $E[|u_i|^2 u_k u_l] = \exp\{j(\theta_k + \theta_l)\} E[|u_i|^2 u_k u_l] \forall \theta_k, \theta_l \in \mathbb{R}$, which holds if and only if $E[|u_i|^2 u_k u_l] = 0$. Similar arguments show that all fourth-order moments necessarily vanish with the exception of $E[|u_i|^4]$ and $E[|u_i|^2 |u_j|^2]$. Now, assume $\mathbf{z} \sim \mathbb{C}N_k(\Sigma)$ with $\Sigma = \mathbf{I}$. Then, by Lemma 3, $r = \|\mathbf{z}\|$ and $\mathbf{u} = \mathbf{z}/r$ are independent, \mathbf{u} has a uniform distribution on $\mathbb{C}S^k$. One can verify from (6), as $g(t) = \exp(-t)$, that the quadratic form $t = r^2 \stackrel{d}{=} 1/2 \chi_{2k}^2$, where χ_{2k}^2 denotes a r.v. from chi-squared distribution with $2k$ degrees of freedom. These facts and the decomposition $\mathbf{z} = r \cdot \mathbf{u}$ yields

$$2 = E[|z_i|^4] = E[t^2] E[|u_i|^4] = \frac{E[(\chi_{2k}^2)^2]}{4} E[|u_i|^4]$$

and

$$1 = E[|z_i|^2 |z_j|^2] = \frac{E[(\chi_{2k}^2)^2]}{4} E[|u_i|^2 |u_j|^2].$$

Then rearranging the terms and using $E[(\chi_{2k}^2)^2] = 4k(k+1)$ gives us the stated claim. [b] This result is a direct consequence of the [a]-part of the Lemma. \blacksquare

Proof of Theorem 2: Using (17) and Theorem 1, the ASC of $\hat{\mathbf{w}}_C$ can be calculated as

$$\begin{aligned} \text{ASC}(\hat{\mathbf{w}}_C; F_\Sigma) &= E[\text{IF}(\mathbf{z}; \mathbf{w}_C, F_\Sigma) \text{IF}(\mathbf{z}; \mathbf{w}_C, F_\Sigma)^H] \\ &= E \left[\left\{ \frac{\alpha_C(r^2)}{\sigma_C} [\mathbf{w} \mathbf{g}^H - \mathbf{I}] \mathbf{b} \right\} \left\{ \frac{\alpha_C(r^2)}{\sigma_C} [\mathbf{w} \mathbf{g}^H - \mathbf{I}] \mathbf{b} \right\}^H \right] \\ &= \frac{E[\alpha_C^2(t)]}{\sigma_C^2} (\mathbf{w} \mathbf{g}^H - \mathbf{I}) E[\mathbf{b} \mathbf{b}^H] (\mathbf{g} \mathbf{w}^H - \mathbf{I}), \quad (25) \end{aligned}$$

where $\mathbf{b} = \mathbf{A}^{-H} \mathbf{u} \mathbf{u}^H \mathbf{A}^H \mathbf{w}$, $r = \|\mathbf{A}^{-1} \mathbf{z}\|$, $t = r^2$, $\mathbf{u} = \mathbf{A}^{-1} \mathbf{z}/r$ and $\Sigma = \mathbf{A} \mathbf{A}^H$. Note that r and \mathbf{u} are independent by Lemma 3 and hence we were able to split the expectation above into two separate parts, other involving r.v. $t = r^2$ and the other involving r.v. \mathbf{u} via \mathbf{b} . Also recall that \mathbf{u} has a uniform

distribution on $\mathbb{C}S^k$ and t has a density (6). Then using (23) we have that

$$\mathbf{b} = \text{vec}(\mathbf{b}) = (\mathbf{w}^T \mathbf{A}^* \otimes \mathbf{A}^{-H}) \mathbf{v} \quad (26)$$

where $\mathbf{v} = \text{vec}(\mathbf{u}\mathbf{u}^H)$. Applying Lemma 4 gives

$$\begin{aligned} E[\mathbf{b}\mathbf{b}^H] &= (\mathbf{w}^T \mathbf{A}^* \otimes \mathbf{A}^{-H}) E[\mathbf{v}\mathbf{v}^H] (\mathbf{w}^T \mathbf{A}^* \otimes \mathbf{A}^{-H})^H \\ &= (\mathbf{w}^T \mathbf{A}^* \otimes \mathbf{A}^{-H}) \left[\frac{\mathbf{I} + \text{vec}(\mathbf{I})\text{vec}(\mathbf{I})^H}{k(k+1)} \right] (\mathbf{w}^T \mathbf{A}^* \otimes \mathbf{A}^{-H})^H \\ &= [k(k+1)]^{-1} [(\mathbf{w}^H \boldsymbol{\Sigma} \mathbf{w})^* \boldsymbol{\Sigma}^{-1} + \mathbf{w}\mathbf{w}^H] \end{aligned}$$

where the last identity follows using (23) and properties of Kronecker product. Next note that

$$(\mathbf{w}^H \boldsymbol{\Sigma} \mathbf{w})^* = \mathbf{w}^H \boldsymbol{\Sigma} \mathbf{w} = (\mathbf{g}^H \boldsymbol{\Sigma}^{-1} \mathbf{g})^{-1}$$

where the first identity follows as $\boldsymbol{\Sigma}$ is positive definite (thus $\mathbf{w}^H \boldsymbol{\Sigma} \mathbf{w} > 0$) and the second identity follows as $\mathbf{w} = \mathbf{W}\mathbf{g}$ and $\mathbf{W}^H \boldsymbol{\Sigma} \mathbf{W} = \boldsymbol{\Sigma}^{-1}$. Hence, $E[\mathbf{b}\mathbf{b}^H] = [k(k+1)]^{-1} [\mathbf{W} + \mathbf{w}\mathbf{w}^H]$ and thus

$$\begin{aligned} &(\mathbf{w}\mathbf{g}^H - \mathbf{I}) E[\mathbf{b}\mathbf{b}^H] (\mathbf{g}\mathbf{w}^H - \mathbf{I}) \\ &= [k(k+1)]^{-1} (\mathbf{w}\mathbf{g}^H - \mathbf{I}) [\mathbf{W} + \mathbf{w}\mathbf{w}^H] (\mathbf{g}\mathbf{w}^H - \mathbf{I}) \\ &= [k(k+1)]^{-1} (\mathbf{w}\mathbf{g}^H - \mathbf{I}) \mathbf{W} (\mathbf{g}\mathbf{w}^H - \mathbf{I}) \\ &= [k(k+1)]^{-1} (\mathbf{W} - \mathbf{w}\mathbf{w}^H). \end{aligned}$$

Here we used the fact that $(\mathbf{w}\mathbf{g}^H - \mathbf{I})\mathbf{w} = \mathbf{0}$ (as $\mathbf{g}^H \mathbf{w} = 1$, which is the constraint imposed by the MVDR beamformer) and the properties $\mathbf{w} = \mathbf{W}\mathbf{g}$ and $\mathbf{W}^H = \mathbf{W}$. Then substituting the previous equation into (25) yields the stated expression for $\text{ASC}(\hat{\mathbf{w}}_{\mathbf{C}}; F_{\boldsymbol{\Sigma}})$.

Similarly, using (17) and Theorem 1, the ASP of $\hat{\mathbf{w}}_{\mathbf{C}}$ can be calculated as

$$\begin{aligned} \text{ASP}(\hat{\mathbf{w}}_{\mathbf{C}}; F_{\boldsymbol{\Sigma}}) &= E[\text{IF}(\mathbf{z}; \mathbf{w}_{\mathbf{C}}, F_{\boldsymbol{\Sigma}}) \text{IF}(\mathbf{z}; \mathbf{w}_{\mathbf{C}}, F_{\boldsymbol{\Sigma}})^T] \\ &= \frac{E[\alpha_{\mathbf{C}}^2(t)]}{\sigma_{\mathbf{C}}^2} (\mathbf{w}\mathbf{g}^H - \mathbf{I}) E[\mathbf{b}\mathbf{b}^T] \\ &\quad \cdot (\mathbf{w}\mathbf{g}^H - \mathbf{I})^T. \end{aligned} \quad (27)$$

Then note that

$$\begin{aligned} E[\mathbf{b}\mathbf{b}^T] &= (\mathbf{w}^T \mathbf{A}^* \otimes \mathbf{A}^{-H}) E[\mathbf{v}\mathbf{v}^T] (\mathbf{A}^H \mathbf{w} \otimes \mathbf{A}^{-*}) \\ &= (\mathbf{w}^T \mathbf{A}^* \otimes \mathbf{A}^{-H}) \left[\frac{\mathbf{I}_{k,k} + \text{vec}(\mathbf{I})\text{vec}(\mathbf{I})^H}{k(k+1)} \right] (\mathbf{A}^H \mathbf{w} \otimes \mathbf{A}^{-*}) \\ &= [k(k+1)]^{-1} \mathbf{w}\mathbf{w}^T [\mathbf{I}_{k,k} + \mathbf{I}] \end{aligned}$$

where we have used Lemma 4 and (26) and rules of calculus involving vec-operator, Kronecker product and commutation matrix listed in (23)–(24). Now it can be inferred that $(\mathbf{w}\mathbf{g}^H - \mathbf{I}) E[\mathbf{b}\mathbf{b}^T] = \mathbf{0}$ (since $(\mathbf{w}\mathbf{g}^H - \mathbf{I})\mathbf{w} = \mathbf{0}$), which subsequently, due to (27), implies that $\text{ASP}(\hat{\mathbf{w}}_{\mathbf{C}}; F_{\boldsymbol{\Sigma}}) = \mathbf{0}$.

ACKNOWLEDGMENT

E. Ollila would like to thank Academy of Finland for supporting this research. The authors would like to thank the anonymous reviewers for their helpful and valuable comments and suggestions.

REFERENCES

- [1] P. R. Krishnaiah and J. Lin, "Complex elliptically symmetric distributions," *Commun. Statist.—Theory Methods*, vol. 15, pp. 3693–3718, 1986.
- [2] R. A. Maronna, "Robust M-estimators of multivariate location and scatter," *Ann. Statist.*, vol. 5, no. 1, pp. 51–67, 1976.
- [3] E. Ollila and V. Koivunen, "Robust antenna array processing using M-estimators of pseudo-covariance," in *Proc. 14th IEEE Int. Symp. Personal, Indoor, Mobile Radio Commun. (PIMRC)*, Beijing, China, Sep. 7–10, 2003, pp. 2659–2663.
- [4] J. Capon, "High resolution frequency-wavenumber spectral analysis," *Proc. IEEE*, vol. 57, pp. 1408–1418, 1969.
- [5] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust Statistics: The Approach Based on Influence Functions*. New York: Wiley, 1986.
- [6] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, pp. 1365–1376, 1987.
- [7] B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 4, pp. 397–401, Jul. 1988.
- [8] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, Jul. 2003.
- [9] A. Gershman, "Robust adaptive beamforming: an overview of recent trends and advances in the field," in *Proc. 4th Int. Conf. Antenna Theory Techniques*, Sep. 9–12, 2003, pp. 30–35.
- [10] R. A. Wooding, "The multivariate distribution of complex normal variables," *Biometrika*, vol. 43, pp. 212–215, 1956.
- [11] N. R. Goodman, "Statistical analysis based on certain multivariate complex Gaussian distribution (an introduction)," *Ann. Math. Statist.*, vol. 34, pp. 152–177, 1963.
- [12] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [13] E. Ollila and V. Koivunen, "Generalized complex elliptical distributions," presented at the 3rd IEEE Sensor Array Multichannel Signal Processing Workshop (SAM), Barcelona, Spain, Jun. 18–21, 2004.
- [14] E. Ollila and V. Koivunen, "Influence functions for array covariance matrix estimators," in *Proc. IEEE Workshop on Statistical Signal Processing (SSP)*, St. Louis, MO, Oct. 1, 2003, pp. 445–448.
- [15] E. Ollila and V. Koivunen, "Robust space-time scatter matrix estimator for broadband antenna arrays," in *Proc. 58th IEEE Vehicular Technology Conference (VTC'03-Fall)*, Orlando, FL, Oct. 6–9, 2003, vol. 1, pp. 55–59.
- [16] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: techniques and key results," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2740–2746, Jun. 2007.
- [17] P. Stoica and R. Moses, *Introduction to Spectral Analysis*. Upper Saddle River, NJ: Prentice-Hall, 1997.
- [18] P. J. Huber, *Robust Statistics*. New York: Wiley, 1981.
- [19] C. Croux, "Limit behavior of the empirical influence function of the median," *Statist. Probab. Lett.*, vol. 37, pp. 331–340, 1998.
- [20] C. Croux and G. Haesbroeck, "Principal component analysis based on robust estimators of the covariance or correlation matrix: Influence functions and efficiencies," *Biometrika*, vol. 87, pp. 603–618, 2000.
- [21] B. Picinbono, "Second order complex random vectors and normal distributions," *IEEE Trans. Signal Process.*, vol. 44, no. 10, pp. 2637–2640, Oct. 1996.
- [22] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.
- [23] A. van den Bos, "The multivariate complex normal distribution—A generalization," *IEEE Trans. Inf. Theory*, vol. 41, no. 2, pp. 537–539, Mar. 1995.

- [24] P. Billingsley, *Probability and Measure*, 3rd ed. New York: Wiley, 1995.
- [25] E. Ollila and V. Koivunen, "Robustness and efficiency analysis of multivariate location and scatter estimators of complex-valued data," Signal Processing Laboratory, Helsinki University of Technology, Espoo, Finland, 02150, 2006.
- [26] J. T. Kent and D. E. Tyler, "Constrained M-estimation for multivariate location and scatter," *Ann. Statist.*, vol. 24, no. 3, pp. 1346–1370, 1996.
- [27] C. D. Richmond, "Pdf's, confidence regions, and relevant statistics for a class of sample covariance-based array processors," *IEEE Trans. Signal Process.*, vol. 44, no. 7, pp. 1779–1793, Jul. 1996.
- [28] J. T. Kent and D. E. Tyler, "Redescending M-estimates of multivariate location and scatter," *Ann. Statist.*, vol. 19, no. 4, pp. 2102–2119, 1991.
- [29] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus*. Chichester, U.K.: Wiley, 1999.



Esa Ollila (M'03) received the M.Sc. degree in mathematics from the University of Oulu, Oulu, Finland, in 1998 and the Ph.D. degree in statistics with honors from the University of Jyväskylä, Finland, in 2002.

From 2004 to 2007, he was a Postdoctoral Fellow of the Academy of Finland at the Signal Processing Laboratory, Helsinki University of Technology, Espoo, Finland. He is currently an Assistant Professor in the Department of Mathematical Sciences, University of Oulu. His research interests focus on statistical signal processing and robust and nonpara-

metric statistical methods.



Visa Koivunen (M'87–SM'98) received the D.Sc. (Tech.) degree with honors from the Department of Electrical Engineering, University of Oulu, Oulu, Finland.

From 1992 to 1995, he was a Visiting Researcher at the University of Pennsylvania, Philadelphia. In 1996, he held a faculty position at the Department of Electrical Engineering, University of Oulu. From August 1997 to August 1999, he was an Associate Professor at the Signal Processing Laboratory, Tampere University of Technology, Finland. Since 1999,

he has been a Professor of signal processing at the Department of Electrical and Communications Engineering, Helsinki University of Technology (HUT), Espoo, Finland. He is one of the Principal Investigators in the SMARAD (Smart Radios and Wireless Systems) Center of Excellence in Radio and Communications Engineering nominated by the Academy of Finland. He has also been Adjunct Full Professor at the University of Pennsylvania, Philadelphia. During his sabbatical leave in 2006–2007, he was Nokia Visiting Fellow at the Nokia Research Center as well as Visiting Fellow at Princeton University, Princeton, NJ. His research interest includes statistical, communications, and sensor array signal processing. He has published more than 260 papers in international scientific conferences and journals.

Dr. Koivunen received the Primus Doctor (best graduate) Award among the doctoral graduates in years 1989 to 1994. He coauthored papers receiving the Best Paper Award in the IEEE PIMRC 2005, EUSIPCO 2006, and EuCAP 2006 conferences. He has been awarded the IEEE Signal Processing Society Best Paper Award for 2007 for a paper coauthored with J. Eriksson. He served as an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS. He is a member of the Editorial Board for the *Signal Processing* journal and the *Journal of Wireless Communication and Networking*. He is also a member of the IEEE Signal Processing for Communication Technical Committee (SPCOM-TC). He was the General Chair of the IEEE Signal Processing Advances in Wireless Communication (SPAWC) 2007 conference in Helsinki, Finland, in June 2007.