A ROBUST SIGNAL SUBSPACE ESTIMATOR

A. Breloy[1], Y. Sun[2], P. Babu[3], G. Ginolhac[4], D.P. Palomar[2], F. Pascal[5]

[1] LEME, Universite Paris Ouest, 50, rue de Sèvres, 92410 Ville d'Avray, France

[2] Department of Electronic and Computer Engineering, HKUST, Clear Water Bay, Kowloon, Hong Kong

[3] CARE, Indian Institute of Technology Delhi, New Delhi-110016, India

[4] LISTIC, Université de Savoie Mont-Blanc, BP 80439, 74944 ANNECY LE VIEUX Cedex, France

[5] L2S, Centrale-Supelec, Plateau du Moulon, 3 rue Joliot-Curie, F-91192 Gif-sur-YvetteCedex, France

ABSTRACT

An original estimator of the orthogonal projector onto the signal subspace is proposed. This estimator is derived as the maximum likelihood estimator for a model of sources plus orthogonal outliers, both with varying power (modeled by Compound Gaussians process), embedded in a white Gaussian noise. Validity and interest - in terms of performance and robustness - of this estimator is illustrated through simulation results on a low rank STAP filtering application.

Index Terms— Subspace Estimation, Robust estimation, Maximum Likelihood, Low Rank, Compound Gaussian, STAP.

1. INTRODUCTION

Subspace estimation is an ubiquitous problem in signal processing [1, 2]. Indeed, in many applications the signal of interest can be accurately modeled as a linear combination of R sources (or modes) embedded in a M (the size of the data) dimensional space . Hence, this signal lies in a subspace of low ($R \ll M$) dimension. The problem of estimating this signal subspace is at the core of numerous methods, such as in DoA estimation [3, 4], interference cancellation [5, 6], blind source separation [7], and dimension reduction.

Classically, the projector onto the signal subspace $\mathbf{\Pi}$ is estimated from the R strongest eigenvectors of the Sample Covariance Matrix (SCM) of the samples set $\{\mathbf{z}_k\}$:

$$\hat{\mathbf{\Sigma}}_{SCM} = \frac{1}{K}\sum_{k=1}^{K}\mathbf{z}_k\mathbf{z}_k^H = \sum_{r=1}^{M}\hat{c}_r\hat{\mathbf{v}}_r\hat{\mathbf{v}}_r^H \ , \tag{1}$$

where $\{\hat{\mathbf{v}}_r\}$ are the SCM eigenvectors and $\{\hat{c}_r\}$ are the corresponding eigenvalues. The projector onto the signal subspace estimated from the SCM is thus denoted:

$$\hat{\mathbf{\Pi}}_{SCM} = \sum_{r=1}^{M}\hat{\mathbf{v}}_r\hat{\mathbf{v}}_r^H = \mathcal{P}_R\left(\hat{\mathbf{\Sigma}}_{SCM}\right) \ , \tag{2}$$

which corresponds to the Maximum Likelihood Estimator (MLE) in the context of deterministic sources embedded in white Gaussian noise (WGN). This estimator yields good results in traditional settings, such as Gaussian noise and high signal to noise ratio (SNR). However, the SCM is known for its poor robustness to heavy-tailed distributed samples, and to the introduction of outliers into the secondary data. Notably, samples corrupted by outliers are distorting the dominant eigensubspace of the SCM (phenomenon also referred to as subspace swap [8]), hence potentially leading to poor performance of an adaptive process built around $\hat{\mathbf{\Pi}}_{SCM}$.

A possible solution to overcome this issue would be to build an estimate as $\hat{\mathbf{\Pi}} = \mathcal{P}_R(\hat{\mathbf{\Sigma}})$, where $\hat{\mathbf{\Sigma}}$ is a covariance matrix (CM) estimator that is robust to various noise distributions and to outliers. To obtain such CM estimators, the current research trend focuses on the robust M-estimators [9,10], defined by a decreasing function φ as:

$$\hat{\mathbf{\Sigma}}_M = \frac{1}{K}\sum_{k=1}^{K}\varphi(\mathbf{z}_k^H\hat{\mathbf{\Sigma}}_M^{-1}\mathbf{z}_k)\mathbf{z}_k\mathbf{z}_k^H \ , \tag{3}$$

that naturally reject outliers through the adaptive weighting of the samples by the factors $\varphi(\mathbf{z}_k^H\hat{\mathbf{\Sigma}}_M^{-1}\mathbf{z}_k)$. Nevertheless, [11,12] exhibited the fact that these CM estimators may reject outliers as well as samples with high SNR (that should rather be promoted in the subspace estimation process). Hence, the M-estimators, may not be the most adapted tool to construct a robust signal subspace estimator.

In [11–15] were developed direct estimators of the signal subspace, that do not rely on an intermediate CM estimate. Especially, [11–14] considered the case of Compound Gaussian [10] sources, leading robust subspace estimators from the point of view of the signal distribution. In this work, we propose to develop a new estimator, that is also robust to the introduction of outliers. To do so, we consider the model of [14] (described and motivated section 2) and we additionally take into account the presence of potential additive outliers (as in [16, 17] and ref. therein). The subspace estimator, derived as the MLE of this model, is shown to be contained in the SVD of a matrix (not a CM estimate) that promotes the samples with high SNR and rejects the ones potentially corrupted by outliers through an adaptive weighting function. Since this MLE is defined as a fixed point in terms of projector, we propose an algorithm based on alternate maximization of the likelihood to compute it. Through a STAP for airborne radar Low-Rank filtering application [5, 6], such solution is shown to reach good performance, both in terms of subspace estimation accuracy and robustness to outliers.

The following convention is adopted: italic indicates a scalar quantity, lower case boldface indicates a vector quantity and upper case boldface a matrix. ^{H} denotes the transpose conjugate operator. $\mathscr{CN}(\mathbf{a}, \boldsymbol{\Sigma})$ is a complex Gaussian vector of mean $\mathbf{a}$ and of covariance matrix $\boldsymbol{\Sigma}$. $\mathbf{I}_M$ is the $M \times M$ identity matrix. $|\boldsymbol{\Sigma}|$ is the determinant of the matrix $\boldsymbol{\Sigma}$. $\{w_n\}$ denotes the set of n elements w_n with $n \in [\![1, N]\!]$.

2. SIGNAL PLUS OUTLIER MODEL

2.1. Model

We assume that K *i.i.d.* samples $\mathbf{z}_k \in \mathbb{C}^M, k \in [\![1, K]\!]$ are available. these samples are modeled as the sum of a signal of interest $\mathbf{c}_k$, a WGN $\mathbf{n}_k$, and a potential outlier $\mathbf{v}_k$:

$$\mathbf{z}_k = \mathbf{c}_k + \mathbf{n}_k + \mathbf{v}_k . \tag{4}$$

• Accounting for possible power fluctuation, the signal $\mathbf{c}_k$ is modeled as a Compound Gaussian (CG) process [10], i.e. a Gaussian random vector with a varying power factor called the *texture*, which is here considered as an unknown deterministic positive parameter for each sample. Therefore, conditionally to the texture τ_k, one has $(\mathbf{c}_k|\tau_k) \sim \mathscr{CN}(\mathbf{0}, \tau_k \boldsymbol{\Sigma}_R)$. The signal is lying in a subspace of dimension $R < M$, assumed to be known, hence its CM is rank deficient and expressed $\boldsymbol{\Sigma}_R = \sum_{r=1}^R c_r \mathbf{v}_r \mathbf{v}_r^H$. As a relaxation hypothesis, we will assume that eigenvalues of $\boldsymbol{\Sigma}_R$ are identical and equal to $c_r = 1 \ \forall r \in [\![1, R]\!]$ (any scaling factor is absorbed by the textures τ_k). This is resulting in a signal CM $\boldsymbol{\Sigma}_R = \boldsymbol{\Pi}_R = \sum_{r=1}^R \mathbf{v}_r \mathbf{v}_r^H$.
• The potential outliers $\mathbf{v}_k$ are also modeled as CG process. Conditionally to their texture β_k one has therefore $(\mathbf{v}_k|\beta_k) \sim \mathscr{CN}(\mathbf{0}, \beta_k \boldsymbol{\Sigma}_o)$. We assume that the outlier CM is also a projection matrix, that is orthogonal to the signal CM : $\boldsymbol{\Sigma}_o = \boldsymbol{\Pi}_R^\perp = \sum_{r=R+1}^M \mathbf{v}_r \mathbf{v}_r^H$.
• The additive WGN is modeled as $\mathbf{n}_k \sim \mathscr{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_M)$, where the variance σ^2 is assumed to be known.

In conclusion, each sample $\mathbf{z}_k$ is distributed conditionally to the textures as:

$$(\mathbf{z}_k|\tau_k, \beta_k) \sim \mathscr{CN}(\mathbf{0}, \boldsymbol{\Sigma}_k) , \tag{5}$$

with

$$\boldsymbol{\Sigma}_k = \tau_k \boldsymbol{\Pi}_R + \beta_k \boldsymbol{\Pi}_R^\perp + \sigma^2 \mathbf{I}_M . \tag{6}$$

The likelihood of the data set, conditioning to $\{\mathbf{v}_r\}$ and the textures parameters, is then:

$$f(\{\mathbf{z}_k\}|\{\mathbf{v}_r\}, \{\tau_k\}, \{\beta_k\}) = \prod_{k=1}^K \frac{e^{-\mathbf{z}_k^H \boldsymbol{\Sigma}_k^{-1} \mathbf{z}_k}}{\pi^M |\boldsymbol{\Sigma}_k|} \tag{7}$$

2.2. Discussion on the hypothesis

We provide in this section the justification of the hypothesis that were made in the previous section.
• Identical eigenvalues of the signal CM: This simplifying

hypothesis is made for obtaining a practical closed-form estimator, as well as deriving a simple algorithm to compute it. Under this hypothesis and without outliers (i.e. $\mathbf{v}_k = \mathbf{0} \ \forall k$) the MLE of the signal subspace projector is given in [14]. In comparison, the exact MLE (assuming non-equals c_r) has no closed form. While different algorithms to compute this exact MLE are proposed in [11, 12], it is shown in [11] that neglecting the difference between eigenvalues has few impact in terms of subspace estimation accuracy. Note that we will test our results on contexts that do not fit the hypothesis $c_r = 1 \ \forall r$ in the simulation section.
• Outlier modeling: Again, this modeling is made for obtaining a closed-form solution with an interesting interpretation. While the proposed modeling can be argued to be seen as a "worst-case contamination", it is acknowledgedly not realistic. Note however that our proposed method will be tested on a more realistic setting, where the outlier is a deterministic vector (e.g. a target) of varying power, so that is spanning only a rank 1 subspace, and which is not necessarily fully orthogonal to the signal subspace.
• Known rank R and WGN variance σ^2: This paper focuses on subspace estimation accuracy. For the sake of simplicity and presentation clarity, we focused on the part of estimating $\boldsymbol{\Pi}_R$, which is the bottleneck of the considered problem. In practice, the proposed algorithm can be called using plug-in estimates of the rank $\hat{R}$ (e.g. [18] and references therein) and WGN power σ^2.

3. MLE OF $\boldsymbol{\Pi}_R$

Theorem 3.1 *The signal subspace orthogonal projector MLE of (7) is given by:*

$$\hat{\boldsymbol{\Pi}}_R = \mathcal{P}_R \left[\hat{\mathbf{R}}(\hat{\boldsymbol{\Pi}}_R) \right] , \tag{8}$$

with

$$\hat{\mathbf{R}}(\hat{\boldsymbol{\Pi}}_R) = \sum_{k=1}^K \rho\left(\hat{\boldsymbol{\Pi}}_R, \mathbf{z}_k\right) \mathbf{z}_k \mathbf{z}_k^H , \tag{9}$$

where the function ρ is defined as

$$\rho\left(\hat{\boldsymbol{\Pi}}_R, \mathbf{z}_k\right) = \frac{max(\ 0\ ,\ \hat{\tau}_k - \hat{\beta}_k\ )}{(\hat{\beta}_k + \sigma^2)\ (\hat{\tau}_k + \sigma^2)} , \tag{10}$$

with the MLE of the textures parameters:

$$\begin{cases} \hat{\tau}_k = max(\ 0\ ,\ \left(\mathbf{z}_k^H \hat{\boldsymbol{\Pi}}_R \mathbf{z}_k\right)/R - \sigma^2) \\ \hat{\beta}_k = max(\ 0\ ,\ \left(\mathbf{z}_k^H \hat{\boldsymbol{\Pi}}_R^\perp \mathbf{z}_k\right)/(M-R) - \sigma^2) \end{cases} . \tag{11}$$

Proof 3.1 *The derivation of this result is essentially identical to the one made in [14]: the MLE of the textures under positivity constraint is expressed and reported in (7) to obtain the reduced log-likelihood. This expression (plus a Lagrange function imposing unitary constraints) is differentiated w.r.t.*

the eigenvectors $\mathbf{v}_r$ defining $\mathbf{\Pi}_R$. *Hence the MLE of $\mathbf{v}_r$'s is shown to be the eigenvectors of $\hat{\mathbf{R}}(\hat{\mathbf{\Pi}}_R)$. A key step is to inherently impose orthogonality between complementary subspaces through the change of variables $\mathbf{\Pi}_R^{\perp} = \mathbf{I} - \mathbf{\Pi}_R$.*

The proposed subspace estimator has an interesting formulation: indeed, it is contained in the dominant eigensubspace of an intermediary matrix that is the SCM of the samples $\mathbf{z}_k$ scaled by adaptive factors $\rho(\hat{\mathbf{\Pi}}_R, \mathbf{z}_k)$. Suppose $\hat{\beta}_k \approx 0$, then $\rho(\hat{\mathbf{\Pi}}_R, \mathbf{z}_k) \approx \hat{\tau}_k/(\sigma^2(\hat{\tau}_k + \sigma^2))$, so the proposed estimator behave exactly as the one of [14], i.e. promotes samples with high SNR in the subspace estimation process. For $\hat{\beta}_k \neq 0$ (suspected presence of an outlier), one can observe two cases. If $\hat{\beta}_k > \hat{\tau}_k$ then the sample is simply rejected. Otherwise, suppose $\sigma^2 \ll \hat{\tau}_k$ and $\beta_k \ll \hat{\tau}_k$, then $\rho(\hat{\mathbf{\Pi}}_R, \mathbf{z}_k) \approx 1/(\sigma^2 + \hat{\beta}_k)$: the contribution of the sample is not canceled but is lowered proportionally to the outlier to noise ratio (ONR). The proposed estimator is therefore susceptible to naturally reject samples that would perturb the subspace estimation process, which is not the case of previous solutions [11, 12, 14].

4. ALGORITHM

Since the proposed estimator is expressed as a fixed point in terms of projector $\hat{\mathbf{\Pi}}_R = \mathcal{P}_R\left[\hat{\mathbf{R}}(\hat{\mathbf{\Pi}}_R)\right]$, one has to resort to iterative algorithms to compute it. As in [14], we propose the following alternate maximization of the likelihood algorithm. Given a starting point $\{\{\tau_k^{(0)}\}, \{\beta_k^{(0)}\}, \mathbf{\Pi}_R^{(0)}\}$, repeat the following two steps until convergence:

• Conditionally to fixed $\mathbf{\Pi}_R^{(n-1)}$ update the textures variables by computing their MLE:

$$\begin{cases} \tau_k^{(n)} = & \max\left(0, \left(\mathbf{z}_k^H \mathbf{\Pi}_R^{(n-1)} \mathbf{z}_k\right)/R - \sigma^2\right) \\ \beta_k^{(n)} = & \max\left(0, \left(\mathbf{z}_k^H \mathbf{\Pi}_R^{\perp (n-1)} \mathbf{z}_k\right)/(M-R) - \sigma^2\right) \end{cases}$$
(12)

• Conditionally to fixed $\{\{\tau_k^{(n)}\}, \{\beta_k^{(n)}\}\}$ update the signal subspace by computing its MLE:

$$\mathbf{\Pi}_R^{(n)} = \mathcal{P}_R\left[\sum_{k=1}^{K} \frac{\max\left(0, \tau_k^{(n)} - \beta_k^{(n)}\right)}{(\beta_k^{(n)} + \sigma^2)(\tau_k^{(n)} + \sigma^2)} \mathbf{z}_k \mathbf{z}_k^H\right]$$
(13)

This process can be encapsulated in the recursion formula:

$$\mathbf{\Pi}_R^{(n+1)} = \mathcal{P}_R\left[\mathbf{R}\left(\mathbf{\Pi}_R^{(n)}\right)\right],$$
(14)

with $\mathbf{R}$ identified from (9), and for which there is currently no proof of convergence in terms of variables. It can however be stated that each update has a unique solution and increases the likelihood. In practice, the iterations are observed to converge and lead to good performance, as illustrated below.

5. APPLICATION TO STAP LOW RANK FILTERING

STAP [19] is applied to airborne radar in order to detect moving targets. Typically, the radar receiver consists in an array of Q antenna elements processing P pulses in a coherent processing interval ($M = PQ$). In the classical framework, one assume that the $K+1$ samples are:

$$\begin{cases} \mathbf{z}_0 = & \alpha_0 \mathbf{d} + \mathbf{c}_0 + \mathbf{n}_0 \\ \mathbf{z}_k = & \mathbf{c}_k + \mathbf{n}_k, \forall k \in [\![1, K]\!] \end{cases},$$
(15)

where the tested cell $\mathbf{z}_0$ contains a target to be filtered, of normalized steering vector $\mathbf{d}$ and SNR α_0/σ^2, radar clutter $\mathbf{c}_0$ and WGN $\mathbf{n}_0$. The samples $\{\mathbf{z}_k\}$ are assumed to be i.i.d. signal-free and outlier-free realizations of the clutter plus noise. These are used to estimate the clutter properties and perform adaptive process on $\mathbf{z}_0$. For this context, the clutter rank $R \ll M$ is known thanks to the Brennan rule [20]. Hence, it is well known that a classical sub-optimal filter (interference canceler) is [5]:

$$\hat{\mathbf{w}}_{lr} = \hat{\mathbf{\Pi}}_c^{\perp} \mathbf{d} = \left(\mathbf{I}_m - \hat{\mathbf{\Pi}}_c\right) \mathbf{d}$$
(16)

where $\hat{\mathbf{\Pi}}_c$ is an estimator of the Clutter Subspace Projector (CSP). The performance of these adaptive filters rely on the accuracy of the estimation of $\mathbf{\Pi}_c$, which therefore illustrates performance of the studied subspace estimators. The considered criterion is the mean SINR-Loss [19]: the ratio between the output SNR of an adaptive filter $\hat{\mathbf{w}}_{lr}$, and the output SNR of the optimal non-adaptive filter $\mathbf{w} = \mathbf{\Sigma}^{-1}\mathbf{d}$, where $\mathbf{\Sigma}$ is the clutter plus noise covariance matrix.

We will compare the performance of the following estimators : $\hat{\mathbf{\Pi}}_{SCM}$, $\hat{\mathbf{\Pi}}_{Ty}$ the estimator build from the SVD of Tyler's M-Estimator [10] (only existing for $K > M$), $\hat{\mathbf{\Pi}}_{MLE}$ the exact MLE of the CSP in the context of CG distributed clutter (computed with algorithm EBMM of [12]), and $\hat{\mathbf{\Pi}}_R$, the proposed robust estimator.

Figure 1 presents the mean SINR-Loss versus the number of secondary data K: One can observe that, in non corrupted context, the proposed estimator reaches performance close to the MLE [12], i.e. the best compared to the state of the art. In Figure 2 the sample $\mathbf{z}_1$ is corrupted by an additive outlier $\mathbf{v}_1 = \alpha_1 \mathbf{d}$ (i.e. proportional to the target to be filtered in $\mathbf{z}_0$). The figure displays the SINR-Loss versus the ONR α_1/σ^2. One can observe that the $\hat{\mathbf{\Pi}}_{SCM}$ and $\hat{\mathbf{\Pi}}_{MLE}$ have performance that quickly drops after a given ONR threshold. $\hat{\mathbf{\Pi}}_{Ty}$ is less impacted by a single corruption, but reaches performance lower than $\hat{\mathbf{\Pi}}_{SCM}$ and $\hat{\mathbf{\Pi}}_{MLE}$ (until ONR= 14dB). $\hat{\mathbf{\Pi}}_R$ offers a good trade-off, as it reaches better performance than $\hat{\mathbf{\Pi}}_{Ty}$, and is more robust than $\hat{\mathbf{\Pi}}_{SCM}$ and $\hat{\mathbf{\Pi}}_{MLE}$ since its performance drops after higher ONR. In Figure 3 multiple samples are corrupted by the same additive outlier $\mathbf{v}_1 = \alpha_1 \mathbf{d}$. The figure displays the SINR-Loss versus the number of samples corrupted for a fixed ONR= 10dB. One can see that the proposed method is highly robust to multiple corruption since its performance do not drop as fast as the ones of other estimators. Hence adaptive filter built from it provides a better interference rejection, even when several training samples are corrupted. In conclusion, the proposed estimator offers the best performance-robustness compromise.
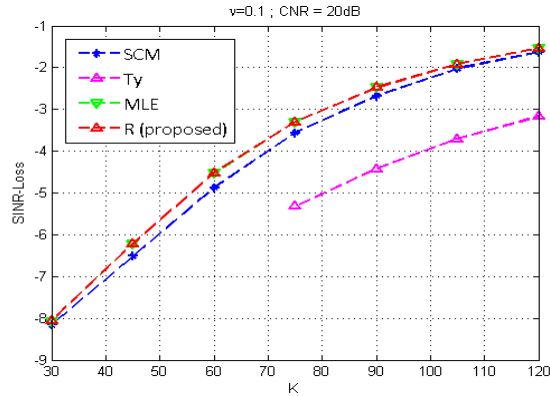
Fig. 1. SINR-Loss versus K. $Q = 8$, $P = 8$, $M = PQ = 64$. Center frequency $f_0 = 450$MHz, Bandwidth $B = 4$MHz. RADAR velocity $100\ m/s$. Inter-element spacing $d = \frac{c}{2f_0}$ (c is the celerity of light). Pulse repetition frequency $f_r = 600$ Hz. $\boldsymbol{\Sigma}_c$ is computed according to the model in [19], its rank is evaluated from the Brennan rule $R = 15$. The CG clutter texture PDF is a Gamma distribution of shape parameter $\nu = 0.1$ and scale parameter $1/\nu$, so the clutter follows a K-distribution. The target $\mathbf{d}$ has a celerity of $V = 35\ m/s$ and is at $+10°$ Azimuth. Clutter to noise ratio is $\text{CNR} = \mathbb{E}(\tau)\text{Tr}(\boldsymbol{\Sigma}_c)/(\sigma^2 R) = 20$dB, and $\sigma^2 = 1$.
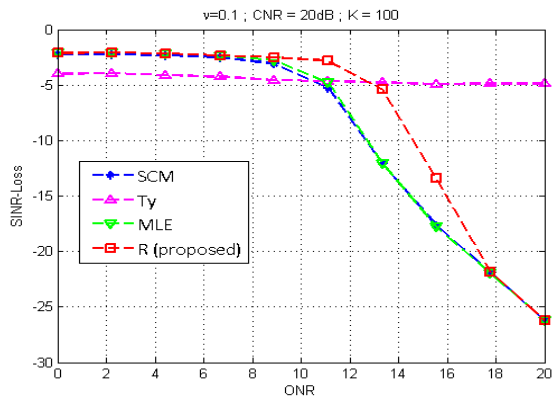

Fig. 2. SINR-Loss versus ONR for one sample corrupted.

Fig. 3. SINR-Loss versus number of corrupted samples. ONR= 10dB

6. REFERENCES

[1] Louis L Scharf, "The svd and reduced rank signal processing," *Signal processing*, vol. 25, no. 2, pp. 113–133, 1991.

[2] A.-J. van der Veen, E.F. Deprettere, and A.L. Swindlehurst, "Subspace-based signal analysis using singular value decomposition," *Proceedings of the IEEE*, vol. 81, no. 9, pp. 1277–1308, Sep 1993.

[3] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans.-ASSP*, vol. 34, no. 3, pp. 276–280, March 1986.

[4] R. Roy and T. Kailath, "ESPRIT-Estimation of signal parameters via rotational invariant techniques," *IEEE Trans.-ASSP*, vol. 37, no. 7, pp. 984–995, July 1989.

[5] I. Kirsteins and D. Tufts, "Adaptive detection using a low rank approximation to a data matrix," *IEEE Trans. on Aero. and Elec. Syst.*, vol. 30, pp. 55 – 67, 1994.

[6] M. Rangaswamy, F.C. Lin, and K.R. Gerlach, "Robust adaptive signal processing methods for heterogeneous radar clutter scenarios," *Signal Processing*, vol. 84, pp. 1653 – 1665, 2004.
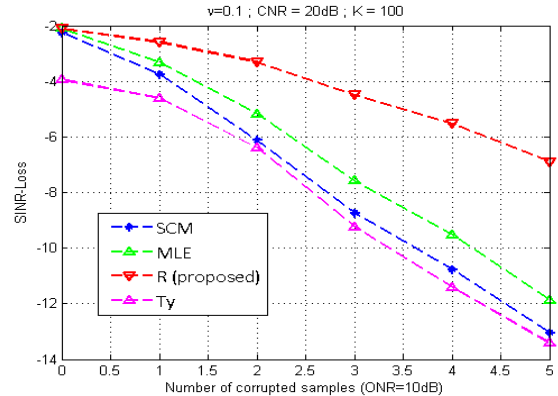
[7] Andrzej Cichocki and Shun-ichi Amari, *Adaptive blind signal and image processing: learning algorithms and applications*, vol. 1, John Wiley & Sons, 2002.

[8] J.K. Thomas, L.L. Scharf, and D.W. Tufts, "The probability of a subspace swap in the svd," *Signal Processing, IEEE Transactions on*, vol. 43, no. 3, pp. 730–736, Mar 1995.

[9] R. A. Maronna, D. R. Martin, and J. V. Yohai, *Robust Statistics: Theory and Methods*, Wiley Series in Probability and Statistics. John Wiley & Sons, 2006.

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

[11] A. Breloy, G. Ginolhac, F. Pascal, and P. Forster, "Clutter subspace estimation in low rank heterogeneous noise context," *Signal Processing, IEEE Transactions on*, vol. 63, no. 9, pp. 2173–2182, May 2015.

[12] Y. Sun, A. Breloy, P. Babu, D.P. Palomar, F. Pascal, and G. Ginolhac, "Low-complexity algorithms for low rank clutter parameters estimation in radar systems," *Signal Processing, IEEE Transactions on, to appear*.

[13] R.S. Raghavan, "Statistical interpretation of a data adaptive clutter subspace estimation algorithm," *IEEE Trans. on Aero. and Elec. Syst.*, vol. 48, no. 2, pp. 1370 – 1384, 2012.

[14] A. Breloy, L. Le Magoarou, G. Ginolhac, F. Pascal, and P. Forster, "Maximum likelihood estimation of clutter subspace in non homogeneous noise context," in *Proceedings of EUSIPCO*, Marrakech, Morocco, September 2013.

[15] O. Besson, N. Dobigeon, and J.-Y. Tourneret, "Minimum mean square distance estimation of a subspace," *Signal Processing, IEEE Transactions on*, vol. 59, no. 12, pp. 5709–5720, Dec 2011.

[16] K. Gerlach, S.D. Blunt, and M.L. Picciolo, "Robust adaptive matched filtering using the fracta algorithm," *IEEE Trans. on Aero.and Elect. Syst.*, vol. 40, no. 3, pp. 929 – 945, July 2004.

[17] O. Besson and S. Bidon, "Adaptive processing with signal contaminated training samples," *Signal Processing, IEEE Transactions on*, vol. 61, no. 17, pp. 4318–4329, Sept 2013.

[18] Petre Stoica and Y. Selen, "Model-order selection: a review of information criterion rules," *Signal Processing Magazine, IEEE*, vol. 21, no. 4, pp. 36–47, July 2004.

[19] J. Ward, "Space-time adaptive processing for airborne radar," Tech. Rep., Lincoln Lab., MIT, Lexington, Mass., USA, December 1994.

[20] L. E. Brennan and F.M. Staudaher, "Subclutter visibility demonstration," Tech. Rep., RL-TR-92-21, Adaptive Sensors Incorporated, March 1992.