Detection of Weak Transient Signals Using a Broadband Subspace Approach

Stephan Weiss[1], Connor Delaosa[1], James Matthews[2], Ian K. Proudler[1], Ben A. Jackson[3]

[1] Department of Electronic & Electrical Engineering, University of Strathclyde, Glasgow G1 1XW, Scotland

[2] PA Consulting, Global Innovation and Technology Centre Cambridge, Melbourn SG8 6DP, UK

[3] UK Defence Science & Technology Laboratory (Dstl), Porton Down, Salisbury SP4 0JQ, UK

{stephan.weiss,connor.delaosa,ian.proudler}@strath.ac.uk, james.matthews@paconsulting.com, bajackson@dstl.gov.uk

Abstract—We investigate the detection of broadband weak transient signals by monitoring a projection of the measurement data onto the noise-only subspace derived from the stationary sources. This projection utilises a broadband subspace decomposition of the data's space-time covariance matrix. The energy in this projected 'syndrome' vector is more discriminative towards the presence or absence of a transient signal than the original data, and can be enhanced by temporal averaging. We investigate the statistics, and indicate in simulations how discrimination can be traded off with the time to reach a decision, as well as with the sample size over which the space-time covariance is estimated.

I. INTRODUCTION

In a number of problems, it is paramount to detect the emergence of a signal whose power may be significantly lower than other sources that are already active in the environment. This is the case e.g. in a cognitive radio scenario [1], where secondary users may be utilising a specific frequency band, but where the emergence of a distant and therefore quiet primary user must be detected in order to instigate the secondary users to vacate this part of the spectrum.

The detection of transient signals is generally based on energy criteria, and can involve a fixed transform such as a wavelet or short-time Fourier transform-type operation to reveal particular patterns of the transient source [2]–[4]. Particularly when multiple measurements are available, data-dependent transforms exploiting the eigenvalue decomposition (EVD) of the covariance matrix of the data can generally attain an optimum compaction of energy into a lower-dimensional subspace in the sense of the Karhunen-Loeve transform [5]; therefore subspace-based methods have emerged that exploit particular partitioning of the space spanned by eigenvectors of the EVD [6]–[10]. While this work has peaked two decades earlier, developments of energy-based subspace detectors are still afoot [11].

The above methods [2]–[4], [6]–[11] operate on narrowband data and calculate an instantaneous covariance matrix that will only capture phase shifts between elements of the data vector. To address the detection of broadband transient signals, it is possible to operate with tapped delay lines or in frequency bins created by a discrete Fourier transform (DFT), where problems

can be treated as narrowband ones. However, when addressing the problems in different bins independently, e.g. for the purpose of parallelisation, generally the coherence between bins is lost, leading to potentially suboptimal solutions [12].

This paper investigates the detection of broadband transient sources via broadband subspace-based methods [13], which are afforded via a space-time covariance matrix that includes an explicit lag parameter and takes both spatial and temporal correlation in the data into account. Similar to narrowband subspace methods, a diagonalisation of this space-time covariance is required to access a subspace decomposition akin to the narrowband case. For the broadband problem, we are looking towards polynomial EVD methods that can decouple the space-time covariance for every lag value [14] — such decompositions have been shown to exist in most case [15], [16] and a number of algorithms have been developed to solves this diagonalisation often with guaranteed convergence [14], [17]–[21].

In the following, we review the broadband signal scenario and the description of the data's second order statistics by a space-time covariance in Sec. II. The broadband subspace decomposition based on the space-time covariance is outlined in Sec. III, leading to the proposed broadband subspace detector in Sec. IV, including an investigation of the subspace energy. The temporal correlation of the latter prohibits the direct description of the subspace energy, but a decorrelation via a decimated processor allows to exploit some known statistics results for generalised χ^2 distributions [22]–[26], and to subsequently define metrics such as the discrimination and decision time in Sec. V. Following some numerical simulations, Sec. VI draws conclusions.

II. WIDEBAND SIGNAL MODEL

A. Source Model

We assume that M sensors acquire a measurement vector $\mathbf{x}[n] \in \mathbb{C}^M$ over discrete time $n \in \mathbb{Z}$, consisting of time series $\mathbf{x}[n] = [x_1[n], \dots, x_M[n]]^{\mathrm{T}} \in \mathbb{C}^M$. This sensor array is illuminated by $L < M$ sources $s_\ell[n]$ $\ell = 1, \dots, L$ via transfer paths with impulse responses $a_{m,\ell}[n]$, $m = 1, \dots, M$, $\ell = 1, \dots, L$, whereby $a_{m,\ell}[n]$ describes the impulse response of the path from the ℓth source to the mth sensor. This path could be a simple delay in the case of a free-space propagation environment, but can equally describe dispersive,
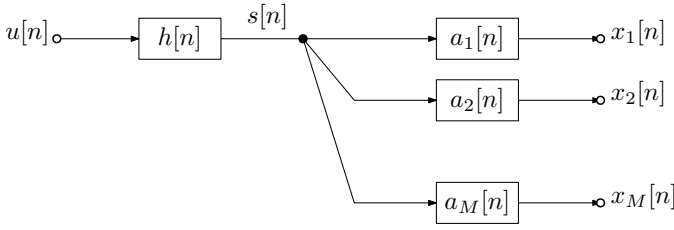
Fig. 1. Source model for a signal signal $s[n]$ contributing to the sensor data. The source is characterised by a zero-mean unit-variance uncorrelated Gaussian noise-excited innovation filter $h[n]$ that generates its desired PSD [27], and a broadband steering vector consisting of transfer functions.

i.e. multipath, channels. For a single source, this scenario is illustrated in Fig. 1; the source signal $s_\ell[n]$ are generated by innovation filters $h_\ell[n]$ excited by zero-mean unit-variance uncorrelated Gaussian noise [27].

If we place the M transfer paths for the ℓth source into a vector

$$\mathbf{a}_\ell[n] = \begin{bmatrix} a_{1,\ell}[n] \\ \vdots \\ a_{m,\ell}[n] \end{bmatrix} , \tag{1}$$

then the data can be modelled as

$$\mathbf{x}[n] = \sum_{\ell=1}^{L} \mathbf{a}_\ell[n] * s_\ell[n] + \mathbf{v}[n] , \tag{2}$$

with additive white zero mean uncorrelated Gaussian noise $\mathbf{v}[n]$, and $*$ denoting the convolution operator. In particular, we assume that all source signals $s_\ell[n]$ as well as the noise in $\mathbf{v}[n]$ are mutually independent. This scenario is depicted in Fig. 1 for the case of a single source in a noise-free environment. For L sources, the data vector $\mathbf{x}[n]$ is obtained by superposition of L models as in Fig. 1.

B. Space-Time Covariance Matrix

To define problems based on cost functions such as the mean square error, we require access to the second order statistics of the data in $\mathbf{x}[n]$. This is captured by the space-time covariance matrix $\mathbf{R}[\tau] = \mathcal{E}\{\mathbf{x}[n]\mathbf{x}^H[n-\tau]\}$, where $\mathcal{E}\{\cdot\}$ is the expectation operator, $\{\cdot\}^H$ a Hermitian transposition, and $\tau \in \mathbb{Z}$ a lag parameter. This matrix contains auto- and cross-correlation sequences, and satisfies the symmetry $\mathbf{R}[\tau] = \mathbf{R}^H[-\tau]$. The z-transform $\mathbf{R}(z) = \sum_\tau \mathbf{R}[\tau]z^{-\tau}$, or in short $\mathbf{R}(z) \bullet\!\!-\!\!\circ \mathbf{R}[\tau]$, is known as the cross-spectral density (CSD) matrix and satisfies the parahermitian property, s.t.. $\mathbf{R}^P(z) = \mathbf{R}^H[1/z^*] = \mathbf{R}(z)$, whereby $\mathbf{R}(z)$ is identical to its parahermitian transpose $\mathbf{R}^P(z)$, which is obtained by Hermitian transposition and time reversal [28].

From (2) and $H_\ell(z) \bullet\!\!-\!\!\circ h_\ell[n]$, and with $\gamma_\ell(z) = H_\ell(z)H_\ell^P(z)$, we can also express the CSD matrix $\mathbf{R}(z)$ as the expansion

$$\mathbf{R}(z) = \sum_{\ell=1}^{L} \mathbf{a}_\ell(z)\mathbf{a}_\ell^P(z)\gamma_\ell(z) + \sigma_v^2\mathbf{I}_M . \tag{3}$$

In (3), a vector $\mathbf{a}_\ell(z)$ is referred to as the broadband steering vector of the ℓth source. A steering vector is used in the beamforming terminology to describe the signature of the source arriving from a specific direction. In the simplest case, this would be a vector of fractional delays [29], [30], but can also be a vector of general, rational transfer functions.

III. Broadband Subspace Decomposition

A. Parahermitian Matrix Eigenvalue Decomposition

If $\mathbf{R}(z)$ arises from a source model such as in Fig. 1 with stable and causal filters $f_\ell[n]$, and transfer paths $a_{m,\ell}[n]$, $\ell = 1,\ldots,L$, $m = 1,\ldots,M$, then $\mathbf{R}(z)$ is analytic. As a result, unless the data vector $\mathbf{x}[n]$ has been multiplexed in the acquisition, the parahermitian matrix $\mathbf{R}(z)$ admits a parahermitian matrix EVD (PhEVD) [15], [16]

$$\mathbf{R}(z) = \mathbf{Q}(z)\mathbf{\Lambda}(z)\mathbf{Q}^P(z) \tag{4}$$

$$= \sum_{m=1}^{M} \mathbf{q}_m(z)\mathbf{q}_m^P(z)\lambda_m(z) . \tag{5}$$

In (4) $\mathbf{Q}(z)$ contains in its columns the eigenvectors $\mathbf{q}_m(z)$, $m = 1,\ldots,M$ of (5). It is a paraunitary matrix such that

$$\mathbf{Q}(z)\mathbf{Q}^P(z) = \mathbf{Q}^P(z)\mathbf{Q}(z) = \mathbf{I}_M , \tag{6}$$

representing a lossless filter bank [28]. The matrix $\mathbf{\Lambda}(z)$ is diagonal and parahermitian, containing the eigenvalues

$$\mathbf{\Lambda}(z) = \text{diag}\{\lambda_1(z),\ \lambda_2(z),\ \cdots,\ \lambda_M(z)\} . \tag{7}$$

Both $\mathbf{Q}(z)$ and $\mathbf{\Lambda}(z)$ are potentially transcendental but analytic functions, i.e. their time domain equivalents converge at least exponentially and can be well-approximated by Laurent polynomials.

B. Uniqueness and Ambiguity

Assuming that the M eigenvalues of $\mathbf{R}(z)$ are distinct and only intersect in a finite number of points, then there is only one solution for the functions $\lambda_m(z)$, $m = 1,\ldots,M$ apart from a permutation. The latter can be addressed by ordering the eigenvalues according to their power, similar how in an ordered EVD eigenvalues are arranged in descending order [31].

For distinct eigenvalues, the associated eigenvectors each exist in uniquely defined 1-d subspaces, but can be multiplied by arbitrary allpass functions. Therefore, w.r.t. the decomposition in (4), the factor $\mathbf{\Lambda}(z)$ is unique, but there is an ambiguity w.r.t. the paraunitary matrix holding the eigenvectors in its columns: if $\mathbf{Q}(z)$ contains valid eigenvectors, then so does $\mathbf{Q}(z)\mathbf{\Phi}(z)$, where $\mathbf{\Phi}(z)$ is a diagonal matrix containing arbitrary allpass filters.

C. Broadband Subspace Partitioning

The PhEVD in (4) admits a subspace decomposition,

$$\mathbf{R}(z) = [\mathbf{U}(z)\ \mathbf{U}_\perp(z)] \begin{bmatrix} \mathbf{\Lambda}_s(z) & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_{\bar{s}}(z) \end{bmatrix} \begin{bmatrix} \mathbf{U}^P(z) \\ \mathbf{U}_\perp^P(z) \end{bmatrix} , \tag{8}$$

where the eigenvalues are grouped into two submatrices $\boldsymbol{\Lambda}_s(z) : \mathbb{C} \to \mathbb{C}^{R \times R}$ and $\boldsymbol{\Lambda}_{\bar{s}}(z) : \mathbb{C} \to \mathbb{C}^{(M-R) \times (M-R)}$. The eigenvectors associated with the eigenvalues in $\boldsymbol{\Lambda}_s(z)$ are the columns of the matrix $\boldsymbol{U}(z) : \mathbb{C} \to \mathbb{C}^{M \times (R}$, which therefore span the subspace within which any components related to these eigenvalues reside. This subspace is orthogonal to its complement, spanned by the columns of $\boldsymbol{U}_\perp(z) : \mathbb{C} \to \mathbb{C}^{M \times (M-R)}$.

A common partitioning of $\boldsymbol{Q}(z)$ in (5) in both the narrowband [31] or the broadband cases [13] is a subspace decomposition that defines signal-plus-noise and noise-only subspaces, where $\boldsymbol{\Lambda}_s(z)$ contains the R signal-related, principal eigenvalues, and $\boldsymbol{U}(z)$ all associated eigenvector components. In contrast, the eigenvalues $\boldsymbol{\Lambda}_{\bar{s}}(z)$ would define the noise floor, and the columns of $\boldsymbol{U}_\perp(z)$ would span the associated noise-only subspace.

IV. Complementary Broadband Matched Subspace Detector

A. Approach

We assume that a number of L sources have been stationary for a period of time, over which a space-time covariance matrix $\hat{\mathbf{R}}[\tau]$ has been estimated, using e.g. the procedures outlined in [32], [33] for the estimation and the optimum support length of this estimate. Using an approximation of the PhEVD in (4) by algorithms of the second order sequential best rotation (SBR2) [14], [17] or sequential matrix diagonalisation (SMD) [18] families to factorise $\hat{\boldsymbol{R}}(z) \bullet\!\!-\!\!\circ \mathbf{R}[\tau]$, we establish the broadband signal-plus-noise subspace spanned by the columns of $\boldsymbol{U}(z)$ and its complement, the noise-only subspace, spanned by the columns of $\boldsymbol{U}_\perp(z)$, as defined in (8) with $R = L$.

If a new source enters the scene, then some of its components may protrude into the noise-only subspace, where a step change in energy can be detected more easily than directly from the measurement $\mathbf{x}[n]$, since the energy contribution of the stationary sources will be removed. We therefore calculate a type of syndrome vector $\mathbf{y}[n] \in \mathbb{C}^{M-L}$ based on $\mathbf{U}_\perp[n] \circ\!\!-\!\!\bullet \boldsymbol{U}_\perp(z)$,

$$\mathbf{y}[n] = \sum_\nu \mathbf{U}_\perp^H[-\nu] \mathbf{x}[n - \nu] . \tag{9}$$

This is a type of projection onto a reduced $(M - K)$-dimensional space, and yields the same energy in $\mathcal{E}\{\|\mathbf{y}[n]\|^2\}$ as a standard projection using the projection operator $\boldsymbol{P}(z) = \boldsymbol{U}_\perp(z)\boldsymbol{U}_\perp^P(z)$, which can be shown via a polynomial singular value decomposition of $\boldsymbol{U}_\perp^P(z)$ [14]. Therefore, the energy of the syndrome $\mathbf{y}[n]$ can be used in a hypothesis test to detect the absence or presence of a transient signal.

B. Signal Statistics

In the absence of a transient signal, let the CSD matrix of $\mathbf{x}[n]$ be $\boldsymbol{R}(z)$ as in (3). Therefore the CSD of the syndrome vector is

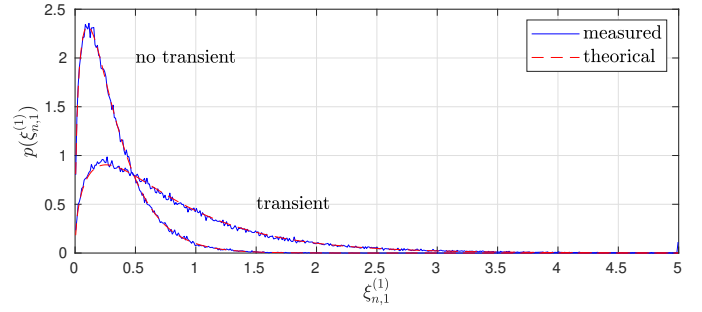$$\boldsymbol{R}_y(z) = \sigma_v^2 \mathbf{I}_{M-L} . \tag{10}$$

Fig. 2. Example of measured and theoretical generalised χ^2 distributions of the energy $\|\mathbf{y}[n]\|_2^2$ in the absence and presence of a transient signal.

If $\mathbf{x}'[n]$ is a modified data vector that in addition to the stationary sources captured in $\mathbf{x}[n]$ includes a transient signal, then a modified CSD matrix $\boldsymbol{R}'(z)$ arises

$$\boldsymbol{R}'(z) = \boldsymbol{R}(z) + \boldsymbol{a}'(z)\gamma'(z)\boldsymbol{a}'^P(z) , \tag{11}$$

and the modified syndrome has the CSD matrix

$$\boldsymbol{R}'_y(z) = \boldsymbol{U}_\perp(z)\boldsymbol{a}'(z)\gamma'(z)\boldsymbol{a}'^P(z)\boldsymbol{U}_\perp^P(z) + \sigma_v^2 \mathbf{I}_{M-L} . \tag{12}$$

The first term on the r.h.s. of (12) creates the offset in power that allows us to potentially detect the presence of a transient signal.

To examine the statistics of $\|\mathbf{y}[n]\|_2^2$, we assume that the noise components of $\mathbf{x}[n]$ are zero-mean and identically distributed Gaussian random variables. However, in the presence of transient components, or subspace leakage of the stationary sources, the elements of the syndrome vector will have Gaussian distributions with different variances. Therefore, $\|\mathbf{y}[n]\|_2^2$ will adhere to a generalised χ^2 distribution [24], [25]. For an example with $M = 6$, $L = 3$, and broadband steering vectors of order 10 drawn from uncorrelated complex Gaussian distributions, Fig. 2 shows the measured distributions in the presence and absence of a transient component in comparison to the theoretical values of a generalised χ^2 distribution based on the implementation in [26].

The discrimination between the cases of a transient source being present or absent can be increase through temporal averaging of energy terms $\|\mathbf{y}[n]\|_2^2$. However, the generalised χ^2 distribution assumes a summation over squared Gaussian distributed, uncorrelated variables. From (12), it is clear that the terms contributed by the transient source will be correlated. Therefore, direct temporal averaging leads to a distribution that will no longer be covered by a generalised χ^2 distribution.

C. Decimated Subspace Detector

In order to deal with large volumes data acquired at a high sampling rate, and to concurrently temporally decorrelate successive syndrome vectors in case of temporal averaging, it is possible to sum over decimated syndrome vectors, such that

$$\xi_{n,D}^{(K)} = \frac{1}{K} \sum_{\mu=0}^{K} \|\mathbf{y}[n - D\mu]\|_2^2 \tag{13}$$
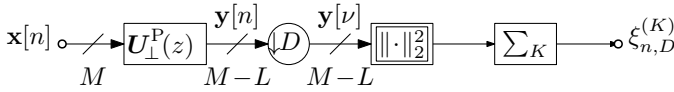
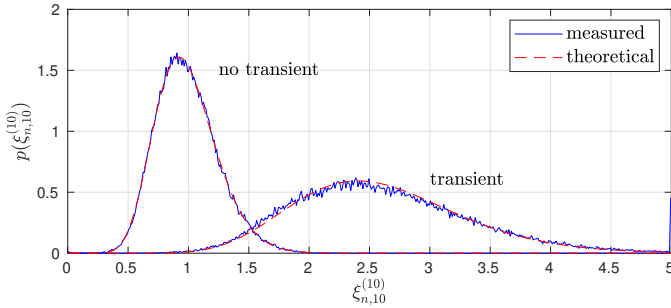Fig. 3. Flow graph of decimated averaging of syndrome energy.

Fig. 4. Example of measured and theoretical generalised χ^2 distributions of the energy $\|\mathbf{y}[n]\|_2^2$ in the absence and presence of a transient signal.

can be used as a criterion for a hypothesis test. The processor is outlined by the flow graph in Fig. 3, where the syndrome energy is decimated by a factor D prior to averaging by a sliding window of length K. It is possible to change the order of the decimator and the processor matrix $\boldsymbol{U}_\perp(z)$ by means of noble identities [28], thus achieving a computationally efficient polyphase implementation of the arrangement in Fig. 3.

The decorrelation, which may be achieved through a decimation by D of the order of the processor matrix $\boldsymbol{U}_\perp(z)$, ensures that $\xi_{n,D}^{(K)}$ follows a generalised χ^2 distribution, which permits theoretical predictions for a hypothesis test on $\xi_{n,D}^{(K)}$. As an example, to highlight both the fit of the distribution but also the increase in discrimination by averaging over a number of syndrome energies, Fig. 4 shows the scenario of Fig. 2 for the case of averaging over $K = 10$ with a decimation by $D = 10$.

V. SIMULATION AND RESULTS

A. Performance Metrics

We define two metrics for the detection of a transient signal. The first is the discrimination of the variable $\xi_{n,D}^{(K)}$ in (13), which can be assessed independently of a specific threshold e.g. by measuring the area under a receiver operating characteristic (ROC) derived from the type of distribution plots in Figs. 2 and 4 [34]. If A is the area under the ROC curve, then its extreme values are $A = \frac{1}{2}$ if the two distributions coincide, and $A = 1$ in case the distributions completely separate. Therefore, we have $\frac{1}{2} \le A \le 1$ and measure the discrimination $\mathcal{D} = 2(1 - A)$, which tends to zero as the distributions increasingly separate. Summing over K terms, and potentially including a decimation by D takes time to evaluate; therefore, a second metric is the decision time $\mathcal{T} = (K - 1)D + 1$, which required to reach the above discrimination of the variable $\xi_{n,D}^{(K)}$. We therefore assess discrimination $\mathcal{D}(\mathcal{T})$ as a function of the decision time $\mathcal{T}$ below.

TABLE I
POWER OF CONTRIBUTIONS FOR REALISTIC CHANNEL SCENARIO.

signal	power
source 1	0.0000 dB
source 2	-4.3494 dB
source 3	-13.2865 dB
noise	-16.2865 dB

B. Test Scenario

Based on multipath propagation responses determine by a radio planning tool in the ISM band with a bandwidth of 20MHz, we have generated transfer functions of order 40 for the case of three sources and $M = 6$ receivers in a dense urban environment. The transfer functions are such that the total power at the receivers for the various sources is as given in Table I. Spatially and temporally uncorrelated additive Gaussian noise corrupts each receiver at 3dB below the weakest source.

We will therefore work with two stationary sources, with the remaining third source — either the medium-powered source 2 or the weakest source 3 — contributing a transient signal. The CSD matrix is either estimated from a finite amount of N samples, or calculated based on the ground truth steering vectors and power spectral densities of the stationary sources, i.e. for $N \longrightarrow \infty$.

C. Discrimination vs Decision Time

Fig. 5 shows two groups of curves. For a first group, sources 1 and 3 from Tab. I make up the stationary signals, and source 2 — of medium power — mimics the transient signal. In this case, the distributions with the transitory source 2 being absent or present separate well, and drop to close to machine accuracy for a only a few number of summation terms K. The estimation of the CSD matrix leads to some subspace leakage [35], but with almost no performance degradation for sample sizes of $N = 10000$ and $N = 1000$, where working with an estimate is almost as good as operating with the ground truth CSD matrix. Still very good discrimination is achieved even when estimating $\hat{\boldsymbol{R}}(z)$ over a very small sample size of $N = 100$. The discrimination time $\mathcal{T}$ here refers to a sampling rate of 20MHz, with a value of $5\mu s$ being equivalent to averaging over $K = 10$ samples at a decimation ratio $D = 10$.

A second group of curves in Fig. 5 refers to sources 1 and 2 forming the group of stationary signals, and the weakest source 3 is taken as the transient signal. In this case, a larger value K for temporal averaging is required to achieve better discrimination, i.e. small values of $\mathcal{D}$. The subspace leakage experienced in the estimation process is now more pronounced for low sample sizes N, such that the approach still works for larger samples sizes but breaks down for $N = 100$, where the low power of the source and the comparatively high noise level lead to no discernible differences in the distributions.
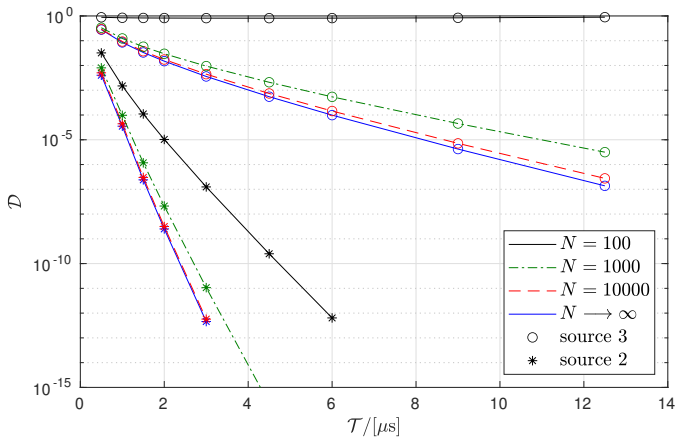
Fig. 5. Plots showing discrimination $\mathcal{D}$ vs decision time $\mathcal{T}$ for detecting sources of different strength, and based on estimates of the space-time covariance relying on various sample sizes N.

VI. CONCLUSION

In this paper we have proposed a broadband subspace-based processor for the detection of potential weak transient signals. By identifying the subspace of stationary sources, we monitor a 'syndrome vector' that gives access to the energy in the noise-only subspace, and can be indicative of emerging source signals. The discrimination of the processor can be enhanced by temporal averaging, whereby a decimation stage reduces the computation complexity as well the correlation in the data, permitting to use generalised χ^2 statistics in assessing the distribution of the data. A simulation has demonstrated that good discrimination is possible, whereby weaker sources can be reliably detected as long as the space-time covariance is estimated with sufficient accuracy, i.e. over a sufficiently long time window.

The evaluation of the discrimination is threshold-independent, and future work would have to focus on setting a suitable detection threshold for a hypothesis test on the absence or presence of a transient signal. Also, the computational burden of the processor rests on the order of the polynomial factorisation of the space-time covariance matrix, with shorter order decompositions emerging [20], [21].

REFERENCES

[1] Z. Quan, S. Cui, H.V. Poor, and A.H. Sayed, "Collaborative wideband sensing for cognitive radios," *IEEE SP Mag.*, **25**(6):60–73, Nov. 2008.

[2] B. Friedlander and B. Porat, "Detection of transient signals by the gabor representation," *IEEE Trans. ASSP*, **37**(2):169–180, Feb. 1989.

[3] ——, "Performance analysis of transient detectors based on a class of linear data transforms," *IEEE Trans. IT*, **38**(2):665–673, Mar. 1992.

[4] B. Porat and B. Friedlander, "Performance analysis of a class of transient detection algorithms-a unified framework," *IEEE Trans. SP*, **40**(10):2536–2546, Oct. 1992.

[5] S. Haykin, *Adaptive Filter Theory*, 2nd ed. Englewood Cliffs: Prentice Hall, 1991.

[6] L.L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. SP*, **42**(8):2146–2157, Aug. 1994.

[7] P. Strobach, "Low rank detection of multichannel gaussian signals using a constrained inverse," in *IEEE ICASSP*, pp. IV/245–IV/248, Apr. 1994.

[8] D. Lundstrom, M. Viberg, and A.M. Zoubir, "Multiple transient estimation using bootstrap and subspace methods," in *9th IEEE SSP*, pp. 184–187, Sep. 1998,

[9] Z. Wang and P. Willett, "A performance study of some transient detectors," *IEEE Trans. SP*, **48**(9):2682–2685, Sep. 2000.

[10] Z. Wang and P. K. Willett, "All-purpose and plug-in power-law detectors for transient signals," *IEEE Trans. SP*, **49**(11):2454–2466, Nov. 2001.

[11] A. Marchioni, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "Subspace energy monitoring for anomaly detection @sensor or @edge," *IEEE Internet of Things Journal*, **7**(8):7575–7589, Aug. 2020.

[12] S. Weiss and I.K. Proudler, "Comparing Efficient Broadband Beamforming Architectures and Their Performance Trade-Offs," in *14th Int. Conf. DSP*, Santorini, Greece, pp. 417–422, July 2002.

[13] M. Alrmah, J. Corr, A. Alzin, K. Thompson, and S. Weiss, "Polynomial subspace decomposition for broadband angle of arrival estimation," in *SSPD*, Edinburgh, Scotland, Sept 2014, pp. 1–5.

[14] J.G. McWhirter, P.D. Baxter, T. Cooper, S. Redif, and J. Foster, "An EVD Algorithm for Para-Hermitian Polynomial Matrices," *IEEE Trans. SP*, **55**(5):2158–2169, May 2007.

[15] S. Weiss, J. Pestana, and I.K. Proudler, "On the existence and uniqueness of the eigenvalue decomposition of a parahermitian matrix," *IEEE Trans. SP*, **66**(10):2659–2672, May 2018.

[16] S. Weiss, J. Pestana, I. Proudler, and F. Coutts, "Corrections to "on the existence and uniqueness of the eigenvalue decomposition of a parahermitian matrix"," *IEEE Trans. SP*, **66**(23):6325–6327, Dec. 2018.

[17] S. Redif, J. McWhirter, and S. Weiss, "Design of FIR paraunitary filter banks for subband coding using a polynomial eigenvalue decomposition," *IEEE Trans. SP*, **59**(11):5253–5264, Nov. 2011.

[18] S. Redif, S. Weiss, and J. McWhirter, "Sequential matrix diagonalization algorithms for polynomial EVD of parahermitian matrices," *IEEE Trans. SP*, **63**(1):81–89, Jan. 2015.

[19] S. Weiss, I. K. Proudler, F. K. Coutts, and J. Pestana, "Iterative approximation of analytic eigenvalues of a parahermitian matrix EVD," in *IEEE ICASSP*, Brighton, UK, May 2019.

[20] S. Weiss, I.K Proudler, F. Coutts, and J. Deeks, "Extraction of analytic eigenvectors from a parahermitian matrix," in *SSPD*, 2020.

[21] S. Weiss, I.K. Proudler, and F.K. Coutts, "Eigenvalue decomposition of a parahermitian matrix: Extraction of analytic eigenvalues," *IEEE Trans. SP*, **69**:722–737, Jan. 2021.

[22] H. Solomon and M.A. Stephens, "Distribution of a sum of weighted chi-square variables," *Journal of the American Statistical Association*, **72**(360):881–885, 1977.

[23] R.B. Davies, "The distribution of a linear combination of χ^2 random variables," *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **29**(3):323–333, 1980.

[24] D. Hammarwall, M. Bengtsson, and B. Ottersten, "Acquiring partial CSI for spatially selective transmission by instantaneous channel norm feedback," *IEEE Trans. SP*, **56**(3):1188–1204, Mar. 2008.

[25] E. Bjornson, D. Hammarwall, and B. Ottersten, "Exploiting quantized channel norm feedback through conditional statistics in arbitrarily correlated mimo systems," *IEEE Trans. SP*, bf 57(10):4027–4041, Oct. 2009.

[26] A. Das and W.S. Geisler, "A method to integrate and classify normal distributions," *arXiv e-prints*, p. arXiv:2012.14331, Dec. 2020.

[27] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed.New York: McGraw-Hill, 1991.

[28] P.P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs: Prentice Hall, 1993.

[29] T.I. Laakso, V. Välimäki, M. Karjalainen, and U.K. Laine, "Splitting the Unit Delay," *IEEE Signal Processing Mag.*, **13**(1):30–60, Jan. 1996.

[30] M. Alrmah, S. Weiss, and J. McWhirter, "Implementation of accurate broadband steering vectors for broadband angle of arrival estimation," in *IET Intelligent Signal Processing*, London, UK, Dec. 2013.

[31] G.H. Golub and C.F. Van Loan, *Matrix Computations*, 3rd ed.Baltimore, Maryland: John Hopkins University Press, 1996.

[32] C. Delaosa, J. Pestana, N.J. Goddard, S. Somasundaram, and S. Weiss, "Sample space-time covariance matrix estimation," in *IEEE ICASSP*, May 2019, pp. 8033–8037.

[33] C. Delaosa, J. Pestana, N.J. Goddard, S. D. Somasundaram, and S. Weiss, "Support estimation of a sample space-time covariance matrix," in *SSPD*, Brighton, UK, Mar. 2019.

[34] J.A. Hanley and B.J. McNeil, "The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve," *Radiology*, vol. 143, pp. 26–36, 1982.

[35] C. Delaosa, F.K. Coutts, J. Pestana, and S. Weiss, "Impact of space-time covariance estimation errors on a parahermitian matrix EVD," in *10th IEEE SAM*, pp. 1–5, July 2018.