Pahalson, C. A. D., Crockett, L. H., & Weiss, S. (in press). Detection of weak transient broadband signals using a polynomial subspace and likelihood ratio test approach. 1-5. Paper presented at 32nd European Signal Processing Conference, Lyon, France.

Detection of Weak Transient Broadband Signals Using a Polynomial Subspace and Likelihood Ratio Test Approach

Cornelius A.D. Pahalson, Louise H. Crockett, and Stephan Weiss
Dept. of Electronic & Electrical Eng., University of Strathclyde, Glasgow, Scotland
{cornelius.pahalson,louise.crockett,stephan.weiss}@strath.ac.uk

***Abstract*—This paper investigates the detection of a weak transient broadband signal. We compare a polynomial subspace detection approach to a likelihood ratio test. While the later is statistically optimal, there may be reasons of computational efficiency and numerical robustness to combine such a likelihood ratio test with the subspace approach. We provide a review of these methods, highlight some of their similarities and differences, and demonstrate in simulations that in particular in the case of detecting very weak signals or signals in very dispersive environments, there may be a compelling argument for combining the likelihood ratio test with a subspace approach.**

I. Introduction

The challenge to fast and reliably detect the emergence of a potentially weak, broadband transient signal is pertinent to a number of applications, such as the detection of primary users in a cognitive radio environment [1]–[3]. In the audio domain, weak speakers may need to be detected in the presence of other speakers [4] or against strong background noise [5]. In a military scenario, there are various uses where the detection of an often weak transient source, either in RF or sonar domains, is desirable if not vital [6].

To detect transient signals, often energy-based criteria are invoked, and utilise transforms such as short-time Fourier transform-type or wavelet-based operations to find transient sources via their correlation structure [7]–[9]. For multi-sensor measurements, data-dependent transforms such as the Karhunen-Loeve transform [10] derived from the covariance matrix of the data via an eigenvalue decomposition (EVD) can attain an optimum energy compaction within a lower-dimensional subspace. Similar subspace partitioning methods have been used in e.g. [11]–[15].

For broadband signals, in [6] a polynomial subspace projection approach has been suggested. Similar to filterbank-based joint source-channel coder in [16], where a projection onto the polynomial nullspace of the code is termed a syndrome and can be used to identify and potentially correct transmission errors, in [6] an increase in energy in the noise-only polynomial subspace is indicative of a new emerging signal in the environment. Here, we want to compare this syndrome approach to the statistically optimum test, the likelihood ratio test (LRT), for Gaussian data [17], [18].

Cornelius Pahalson acknowledges funding support from XYZ.

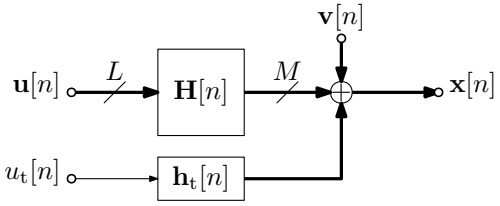

Fig. 1. Signal model for measurement $\mathbf{x}[n] \in \mathbb{C}^M$.

Below, based on the system model outlined in Sec. II, Sec. III reviews the syndrome approach and Sec. IV the LRT method. We suggest a combination of the two by applying the LRT to the polynomial subspace-projected data, which we evaluate in Sec. V.

II. Signal Model and Space-Time Covariance

A. Signal Model

We assume a scenario with M sensors measuring emissions from an environment containing $L < M$ stationary sources, as shown in Fig. 1. For this scenario, the sources are modelled as mutually independent, temporally uncorrelated zero mean and unit variance signals $u_\ell[n]$, $\ell = 1, \ldots, L$ gathered in a vector $\mathbf{u}[n] = [u_1[n], \ldots, u_L[n]]^{\mathrm{T}}$. Both the power spectral density of the individual sources as well as a convolutive mixing is performed by a matrix of impulse responses $\mathbf{H}[n] \in \mathbb{C}^{M\times L}$. Further, the measurements are corrupted by additive Gaussian noise $\mathbf{v}[n] \in \mathbb{C}$ with covariance $\mathcal{E}\{\mathbf{v}[n]\mathbf{v}[n-\tau]\} = \sigma_v^2\delta[\tau]\mathbf{I}$. A further $(L+1)$st signal illuminates the array, which is tied via a vector of filters $\mathbf{h}_{\mathrm{t}}[n] \in \mathbb{C}^M$ to another zero mean unit variance uncorrelated Gaussian signal $u_{L+1}[n]$. It is the presence of this signal that we would like to detect.

B. Space-Time Covariance Matrices

For the space-time covariance $\mathbf{R}[\tau] = \mathcal{E}\{\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n]\}$ of the measurement vector $\mathbf{x}[n]$, we first consider the case $u_{L+1}[n] = 0$. We can state the z-transform of the space-time covariance, the cross-spectral density $\boldsymbol{R}(z) = \sum_\tau \mathbf{R}[\tau]z^{-\tau}$, or short $\mathbf{R}[\tau] \circ\!\!-\!\!\bullet\, \boldsymbol{R}(z)$, as

$$\boldsymbol{R}(z) = \boldsymbol{H}(z)\boldsymbol{H}^{\mathrm{P}}(z) + \sigma_v^2\mathbf{I}\,. \tag{1}$$

In (1), we have $\boldsymbol{H}(z) \bullet\!\!-\!\!\circ\, \mathbf{H}[n]$ as a matrix of transfer functions; the parahermitian operator $\{\cdot\}^{\mathrm{P}}$ implements a Hermitian transposition and time reversal, such that $\boldsymbol{H}^{\mathrm{P}}(z) =$

$\{\boldsymbol{H}(1/z^*)\}^{\mathrm{H}}$ [19]. As a result, $\boldsymbol{R}(z)$ is a parahermitian matrix, i.e. it satisfies $\boldsymbol{R}^{\mathrm{P}}(z) = \boldsymbol{R}(z)$.

In the case of $u_{L+1}[n] \neq 0$, its contribution to the overall CSD is

$$\boldsymbol{R}_{\mathrm{t}}(z) = \boldsymbol{h}_{\mathrm{t}}(z)\boldsymbol{h}^{\mathrm{P}}{}_{\mathrm{t}}(z) \;, \tag{2}$$

where $\boldsymbol{h}_{\mathrm{t}}(z) \bullet\!\!-\!\!\circ\, \boldsymbol{h}[n]$. Note that $\boldsymbol{R}_{\mathrm{t}}(z)$ has rank one only, and that for the overall CSD we have $\boldsymbol{R}(z) + \boldsymbol{R}_{\mathrm{t}}(z)$.

III. Polynomial Subspace Detection

A. Analytic Eigenvalue Decomposition

The parahermitian nature of $\boldsymbol{R}(z)$ generally admits an analytic EVD [20]–[22]

$$\boldsymbol{R}(z) = \boldsymbol{Q}(z)\boldsymbol{\Lambda}(z)\boldsymbol{Q}^{\mathrm{P}}(z) \;, \tag{3}$$

where $\boldsymbol{\Lambda}(z)$ is a diagonal parahermitian matrix containing the eigenvalues of $\boldsymbol{R}(z)$, and $\boldsymbol{Q}(z)$ is paraunitary, such that $\boldsymbol{Q}^{-1}(z) = \boldsymbol{Q}(z)$, holding the corresponding eigenvectors in its columns. Both $\boldsymbol{\Lambda}(z)$ and $\boldsymbol{Q}(z)$ exist as analytic functions; this is an important property, as it allows to approximate particularly $\boldsymbol{Q}(z)$ arbitrarily closely by polynomials of sufficient order simply by shifts and truncations [23].

Given the source model in Fig. 1, it is possible to utilise (3) in order to define a subspace decomposition,

$$\boldsymbol{\Lambda}(z) = \mathrm{blockdiag}\left\{\boldsymbol{\Lambda}_H(z) + \sigma_v^2\mathbf{I}_L,\;\; \sigma_v^2\mathbf{I}_{M-L}\right\} \;, \tag{4}$$

$$\boldsymbol{Q}(z) = [\boldsymbol{Q}_{\parallel}(z),\;\; \boldsymbol{Q}_{\perp}(z)] \;, \tag{5}$$

where $\boldsymbol{\Lambda}_H(z)$ contains the L non-zero analytic eigenvalues of $\boldsymbol{H}(z)\boldsymbol{H}^{\mathrm{P}}(z) : \mathbb{C} \to \mathbb{C}^{L\times L}$, and $\boldsymbol{Q}_{\parallel}(z)$ holds the corresponding eigenvectors. The remaining analytic eigenvectors in $\boldsymbol{Q}_{\perp}(z)$ span the noise-only subspace of $\boldsymbol{R}(z)$, which is not reached by any of the L sources.

B. Subspace Projection and Syndrome Vector

The subspace decomposition afforded by $\boldsymbol{Q}_{\perp}(z)$ has been used to detect an emerging weak broadband signal $u_{L+1}[n]$ in [6] by projecting $\mathbf{x}[n]$ onto the noise-only subspace via

$$\mathbf{s}[n] = \sum_{\nu} \mathbf{Q}_{\perp}^{\mathrm{H}}[-\nu]\mathbf{x}[n-\nu] \;, \tag{6}$$

as also depicted in 2. The resulting $\mathbf{s}[n] \in \mathbb{C}^{M-L}$ has been referred to as a syndrome vector, as in the absence of $u_{L+1}[n]$, we have $\mathbf{s}[n] \in \mathcal{N}(\underline{0}, \sigma_v^2\mathbf{I}_{M-L})$. In case of $u_{L+1}[n] \neq 0$, some of its energy will leak into the noise-only subspace, where it will increase the variance of the syndrome and thus may be detected. To some extend, this is a generalisation of narrowband subspace detection approaches [11], and its broadband extension via (6) has found use for e.g. voice activity detection in the presence of stronger speakers [4] or noise [5].

If temporal correlation in $\mathbf{s}[n]$ is ignored, then a detection variable $\|\mathbf{s}[n]\|_2^2$ will follow a generalised χ^2 distribution. If additional temporal averaging is applied, the temporal correlation limits the performance. This temporal decorrelation can be somewhat broken up by averaging over a decimated output $\mathbf{s}[nQ]$ [6], where Q exceeds the period over which $\mathbf{s}[n]$ is correlated.

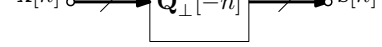

Fig. 2. Projection of measurements $\mathbf{x}[n]$ onto the noise-only subspace, and syndrome vector $\mathbf{s}[n]$.

IV. Likelihood Ratio Test Application

A. Likelihood Ratio Test Considerations

In order to compare or to combine the subspace test with a likelihood ratio test, we explore the latter for a general decision on a vector $\mathbf{y}_n \in \mathbb{C}^K$, which can later be constructed from measurement vectors $\mathbf{x}[n]$ or syndrome vectors $\mathbf{s}[n]$. The dimension K will depend on this choice. This vector $\mathbf{y}_n$ can have two independent components: $\mathbf{y}_{0,n} \in \mathbb{C}^K$, which is the ever-present stationary noise, and $\mathbf{y}_{1,n} \in \mathbb{C}^K$, which represents the transient component. These signals are assumed to be zero mean with covariance matrices $\mathbf{R}_0$ and $\mathbf{R}_1$, respectively. We want to distinguish the two hypotheses

$$\begin{aligned} H_0 :&\quad \mathbf{y}_n = \mathbf{y}_{0,n} \;, \\ H_1 :&\quad \mathbf{y}_n = \mathbf{y}_{0,n} + \mathbf{y}_{1,n} \;, \end{aligned}$$

where $\mathbf{y}_{0,n}$ is the contribution from the L stationary sources and additive noise, and $\mathbf{y}_{1,n}$ is due to the transient signal. For zero-mean Gaussian signals $\mathbf{y}_i \sim \mathcal{N}(\underline{0}, \mathbf{R}_i)$, $i = 0, 1$, the probability density functions of $\mathbf{y}$ for the two cases are

$$p(\mathbf{y}_n|H_0) = (2\pi|\mathbf{R}_0|)^{-\frac{1}{2}}\mathrm{e}^{-\frac{1}{2}\mathbf{y}_n^{\mathrm{H}}\mathbf{R}_0^{-1}\mathbf{y}_n} \;, \tag{7}$$

$$p(\mathbf{y}_n|H_1) = (2\pi|\mathbf{R}_0 + \mathbf{R}_1|)^{-\frac{1}{2}}\mathrm{e}^{-\frac{1}{2}\mathbf{y}_n^{\mathrm{H}}(\mathbf{R}_0+\mathbf{R}_1)^{-1}\mathbf{y}_n} \;, \tag{8}$$

with $|\cdot|$ denoting the determinant of its matrix argument.

For the likelihood ratio $L(\mathbf{y}_n)$, we have

$$L(\mathbf{y}_n) = \frac{p(\mathbf{y}_n|H_0)}{p(\mathbf{y}_n|H_1)} = \frac{|\mathbf{R}_0 + \mathbf{R}_1|^{\frac{1}{2}}}{|\mathbf{R}_0|^{\frac{1}{2}}}\mathrm{e}^{-\frac{1}{2}\mathbf{y}_n^{\mathrm{H}}\mathbf{A}\mathbf{y}_n} \;, \tag{9}$$

with

$$\mathbf{A} = \mathbf{R}_0^{-1} - (\mathbf{R}_0 + \mathbf{R}_1)^{-1} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{\mathrm{H}} \;. \tag{10}$$

Thus,

$$L(\mathbf{y}_n) = \frac{|\mathbf{R}_0 + \mathbf{R}_1|^{\frac{1}{2}}}{|\mathbf{R}_0|^{\frac{1}{2}}}\mathrm{e}^{-\frac{1}{2}\|\boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{Q}^{\mathrm{H}}\mathbf{y}_n\|_2^2} \;, \tag{11}$$

represents the likelihood ratio.

We now need to find a threshold c to accept or reject the hypothesis,

$$L(\mathbf{y}_n) \underset{H_1}{\overset{H_0}{\lessgtr}} c \;. \tag{12}$$

This leads to

$$\|\boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{Q}^{\mathrm{H}}\mathbf{y}_n\| \underset{H_0}{\overset{H_1}{\gtrless}} 2\ln\left\{\frac{|\mathbf{R}_0|^{\frac{1}{2}}}{|\mathbf{R}_0 + \mathbf{R}_1|^{\frac{1}{2}}}c\right\} = c' \;. \tag{13}$$

Thus, the term $\|\boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{Q}^{\mathrm{H}}\mathbf{y}_n\|$ defines the test statistic.

B. LRT Applied to Measurements

In the case that we want to work with a window of T snapshots of the measurement vector $\mathbf{x}[n]$, our variable for the LRT is

$$\mathbf{y}_n^{\mathrm{H}} = [\mathbf{x}_n^{\mathrm{H}}, \mathbf{x}_{n-1}^{\mathrm{H}}, \ldots, \mathbf{x}_{n-T+1}^{\mathrm{H}}] \, . \tag{14}$$

With the space time covariances defined in Sec. II-B, we have in this case

$$\mathbf{R}_0 = \begin{bmatrix} \mathbf{R}[0] & \ldots & \mathbf{R}[T-1] \\ \vdots & \ddots & \vdots \\ \mathbf{R}[1-T] & \ldots & \mathbf{R}[0] \end{bmatrix} \, . \tag{15}$$

Likewise, $\mathbf{R}_1$ can be constructed from lag components of $\mathbf{R}_{\mathrm{t}}[\tau]$. While we know that $\mathbf{R}_{\mathrm{t}}[\tau]$ is rank one, we can only say that $\mathbf{R}_1$ will a most possess rank T [24].

In terms of dimensions, we have $\mathbf{R}_i \in \mathbb{C}^{(MT)\times(MT)}$, $i = 0, 1$. Thus, for large values of M and T, the inversion can be computationally and numerically challenging. Typically in practice, in the presence of the transient signal, we measure $\mathbf{R}_0 + \mathbf{R}_1$. If the covariances are available separately, then once for (10) the inverse $\mathbf{R}_0^{-1}$ has been calculated, we can determine the low rank update $(\mathbf{R}_0 + \mathbf{R}_1)^{-1}$ via the matrix inversion lemma [25].

C. LRT Applied to Syndrome

If the LRT is applied to T snapshots of the syndrome vector $\mathbf{y}[n]$, we work with a variable

$$\mathbf{y}_n^{\mathrm{H}} = [\mathbf{s}_n^{\mathrm{H}}, \mathbf{s}_{n-1}^{\mathrm{H}}, \ldots, \mathbf{s}_{n-T+1}^{\mathrm{H}}] \, . \tag{16}$$

Taking note of its space-time covariance $\mathbf{R}'[\tau] = \mathcal{E}\{\mathbf{y}[n]\mathbf{y}^{\mathrm{H}}[n-\tau]\}$ and its corresponding CSD matrix $\boldsymbol{R}'(z) : \mathbb{C} \to \mathbb{C}^{(M-L)\times(M-l)}$,

$$\boldsymbol{R}'(z) = \boldsymbol{Q}^{\mathrm{P}}{}_{\perp}(z)\boldsymbol{R}(z)\boldsymbol{Q}_{\perp}(z) \, , \tag{17}$$

we can construct $\mathbf{R}_0$ akin to (15). Similarly, $\mathbf{R}_1$ can be obtained. We now have covariance matrices $\mathbf{R}_i$, $i = 0, 1$ of size $T(M-L) \times T(M-L)$.

With respect to the LRT applied to measurement data in Sec. IV-B, we disregard any possible distinction between the stationary sources and the transient signal in the signal-plus-noise subspace of $\boldsymbol{R}(z)$. However, if the transient source is weak, two advantages arise: firstly, by suppressing the dominant eigenvalues of $\mathbf{R}(z)$ in $\mathbf{R}'$, the condition number of the matrices $\mathbf{R}_i$ will be reduced for the LRT applied to the syndrome versus the one applied to the measurement data. Secondly, the covariance matrices are now of a reduced size. Thus, the matrix inversions become numerically and computationally less challenging.

With respect to just assessing the energy on the syndrome vector as in Sec. III-B, there the evaluation of $\|\mathbf{s}[n]\|^2$ treats all components of the noise-only subspace equal, while the LRT weights contributions by the term $\boldsymbol{\Lambda}^{\frac{1}{2}}$; in the case of the LRT applied to the syndrome vector, this will reflect the eigenvalues of the transient source (offset by the variance σ_v^2 of the additive noise) within the noise-only subspace of the stationary sources. Hence, subspace components with a stronger contribution by the transient signal will be weighted stronger, thus emphasising the difference between the more heuristic syndrome energy approach in [6] and the statistically optimum LRT.

V. Simulations and Results

We now want to compare the three tests: the simple syndrome approach of Sec. III-B, the LRT applied to the measurement data in Sec. IV-B, and the proposed LRT based on the syndrome in Sec. IV-C.

A. Performance Metrics, Algorithms, and Parameters

We generally want to assess the separation of distributions when data belongs to either of the two hypotheses. A useful metric is the receiver operating characteristic (ROC) [26], which in our case evaluates the probability of correctly detecting the transient signal versus the false alarm of incorrect detecting a transient signal, i.e. with respect to the test in (13) the threshold c' is exceeded in the absence of any transient signal.

To obtain a single numerical value assessing how well two distributions separate, we also define the separation distance δ as

$$\delta = \frac{|\mu_1 - \mu_0|}{(\sigma_0 + \sigma_1)/2} \, . \tag{18}$$

This ratio assesses the difference of the means μ_i , normalised by the mean of the standard deviations σ_i, for the two hypotheses H_i, $i = 0, 1$.

For the general test setup, w.r.t. the model in Fig. 1 we assume $M = 8$ sensors and $L = 5$ stationary sources. When present, the transient signal sits in the noise floor, and possesses a power equal to the variance σ_v^2 of the additive noise term $\mathbf{v}[n]$ in Fig. 1. We will alter the signal to noise ratio (SNR) γ measured between the stationary sources and the additive noise (and therefore the transient source) as

$$\gamma = \frac{\mathrm{tr}\{\mathbf{R}[0]\} - M\sigma_v^2}{M\sigma_v^2} \, , \tag{19}$$

where $\mathrm{tr}\{\cdot\}$ is the trace operator. Other variable parameters in the tests are the temporal window T, and the dispersiveness of the mixing systems, measured by the time domain support of $\mathbf{H}[n]$ and $\mathbf{h}_{\mathrm{t}}[n]$ in Fig. 1, which equals the length K of the finite impulse response filters of these systems.

In term of algorithms, we assess the LRT applied to the measurements and to the syndrome, labelled LRT($\mathbf{x}$) and LRT($\mathbf{s}$) respectively. Since the syndrome approach in [6] ignores temporal correlation, as a reference we include an LRT approach that only utilises $\mathbf{R}[0]$ and $\mathbf{R}_{\mathrm{t}}[0]$ for its construction, labelled LRT$_2(\mathbf{x})$, i.e. it also ignores the temporal correlation of the data. We also show results with space-time covariance matrices estimated over 10^5 samples using a best linear unbiased estimator [27], which essentially provides the performance of a generalised LRT [18].

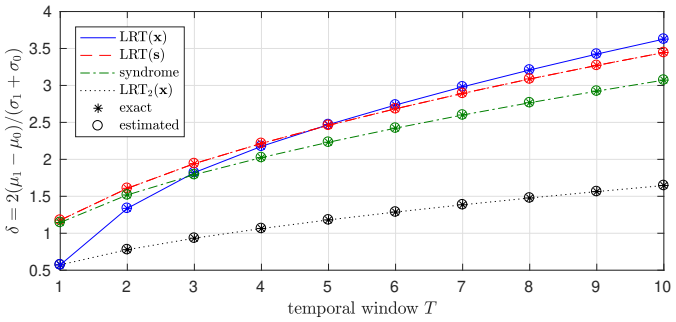

Fig. 3. Separability of distributions for different approaches with $M = 8$, $L = 5$, $K = 8$, and an SNR $\gamma = 10$dB.

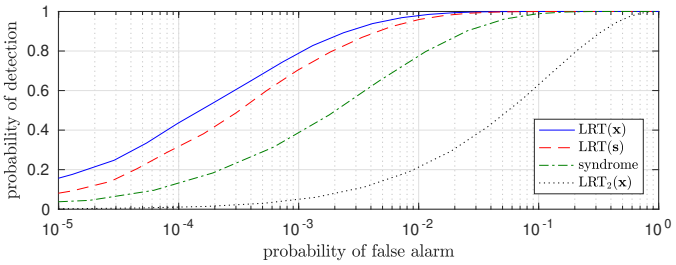

Fig. 4. ROC curve for the case $T = 10$ in Fig. 3.

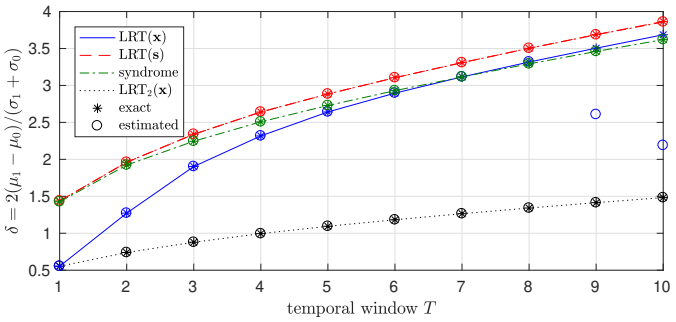

Fig. 5. Separability with $K = 16$ compared to Fig. 3.

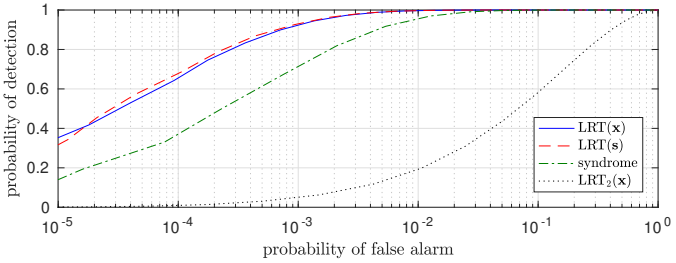

Fig. 6. ROC curve for the case $T = 10$ in Fig. 5.

B. Basic Simulation

With the fixed parameters as specified in Sec. V-A, Figs. 3 and 4 show the simulation results for a mixing system with $K = 8$ and an SNR of $\gamma = 10$dB, i.e. the stronger stationary sources are 10dB above the noise and the transient signal. In Fig. 3, the separability δ as defined in (18) is given for different temporal windows $T = 1, \ldots, 10$. For $T = 1$, there is temporal correlation to consider, and LRT($\mathbf{x}$) and LRT$_2$($\mathbf{x}$) perform identical. As T increases, both methods include more averaging, and the separability increases; taking the correct temporal correlation into account provides a higher increase in δ for LRT($\mathbf{x}$) than for LRT$_2$($\mathbf{x}$).

In comparison, the syndrome-based methods start with a higher separability for $T = 1$ than the tests operating on $\mathbf{x}$. Since LRT($\mathbf{x}$) is the statistically optimum approach, this is suprising and likely due to the temporal decorrelation w.r.t. the stationary sources applied by $\boldsymbol{Q}_\perp(z)$. Eventually, as T is increased, the resulting performance advantage of LRT($\mathbf{s}$) w.r.t. LRT($\mathbf{x}$) evaporates, and the latter shows the best performance for $T > 5$ in this situation. Interestingly, the syndrome approach [6], which different from LRT($\mathbf{s}$) does not apply any weighting of the contributions in the noise-only subspace and ignores temporal correlation, performs reasonably well and much better then LRT$_2$($\mathbf{x}$) — it is again the decorrelating property of $\boldsymbol{Q}_\perp(z)$ that may provide a benefit.

Since δ is a simplified assessment of how well distributions separate, Fig. 4 shows the ROC curve for $T = 10$. Here, there is an advantage for the LRT approach to assess the entirety of the available data in the measurements $\mathbf{x}[n]$ rather than in the preprocessed syndrome $\mathbf{s}[n]$. Note that in both Figs. 3 and 4, the estimates are sufficiently accurate to not cause any drop in performance.

C. Changing SNR and Support of Mixing System

In order to harden the suspicion that in the presence of temporal correlation it is advantageous to work with the syndrome vector preprocessed by the $\boldsymbol{Q}_\perp(z)$, we increase the temporal correlation by doubling the support of $\boldsymbol{H}(z)$ to $K = 16$. The result on the separability in Fig. 5 indeed shows that the performance gap widens compared to Sec. V-B, and that over the range of $0 < T \leq 10$, the LRT directly applied to the measurement data does not anymore reach the performance of the syndrome-based LRT. This is also highlighted by the ROC curve for $T = 10$ in Fig. 6.

In the previous simulation, the difference in performance between LRT based on precise knowledge and GLRT based on estimated covariance matrices was neglible. In Fig. 5, there is a break-down in performance for GLRT based on the measurement data for $T > 8$. Here, in the case of $T = 10$, the matrix $\mathbf{R}_0$ for LRT($\mathbf{s}$) is 30×30, while for LRT($\mathbf{x}$) it is 80×80 with a condition number that is increased by 3 orders of magnitude w.r.t. the sybdrome-based case. Hence, any estimation errors tend to have a more severe impact and may lead to very poor results.

In addition to increasing K w.r.t. the experiment in Sec. V-B, we now also increase the power of the stationary sources from $\gamma = 10$dB to $\gamma = 20$dB. This has the effect of increasing the condition number of the covariance matrices when based on $\mathbf{x}[n]$ rather than $\mathbf{s}[n]$. It also means that it becomes more difficult for LRT($\mathbf{x}$) to extract information on the transient source from the signal subspace of the now even stronger stationary sources. Considering the separability results in Fig. 7, the reason that it cannot perform as well as LRT($\mathbf{s}$) is likely due to the decorrelating nature of $\boldsymbol{Q}_\perp(z)$. Both effects together increase the advantage of the proposed LRT($\mathbf{s}$) over LRT($\mathbf{x}$). Note that while the simplistic measure

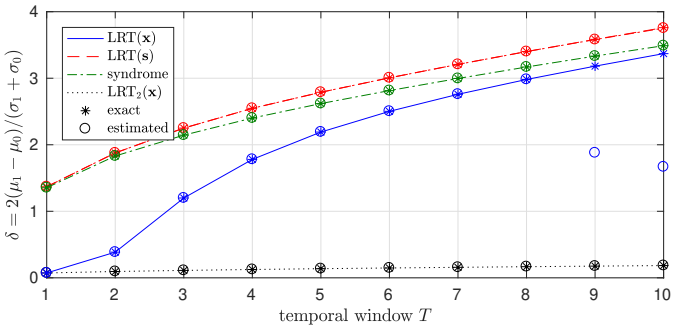

Fig. 7. Separability with $K = 16$ and SNR $\gamma = 20$dB compared to Fig. 3.

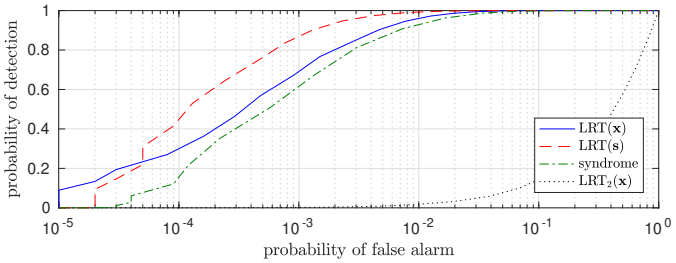

Fig. 8. ROC curve for the case $T = 10$ in Fig. 7.

δ suggests a marginally better separation for the syndrome approach in [6] than for LRT(x) at $T = 10$, considering the ROC curve in Fig. 8 reveals somewhat reversed fortunates when considering the actual distributions. This, though, does not affect the superior performance of LRT(s) over LRT(x)

VI. Conclusions

We have considered the detection of weak transient signals using both a polynomial subspace projection approach and a likelihood ratio test. Both can be derived from the space-time covariance matrix of the data. We have shown that the consideration of temporal as well as spatial correlations is vital to the performance of the detector; this is not surprising. As a remarkable result though, simulation results indicate that both for detecting very weak signals, as well as for operating in very dispersive mixing environments, there is an advantage to restrict the likelihood ratio test to the projected (here referred to as syndrome) data rather than applying it directly to the measurements. This is likely due to the decorrelating effect that the subspace projection operation has on the data.

References

[1] Z. Quan, S. Cui, H. V. Poor, and A. H. Sayed, "Collaborative wideband sensing for cognitive radios," *IEEE Signal Processing Magazine*, vol. 25, no. 6, pp. 60–73, Nov. 2008.

[2] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio : State-of-the-art and recent advances," *IEEE Signal Processing Magazine*, vol. 29, no. 3, pp. 101–116, May 2012.

[3] M. H. Al-Ali and K. C. Ho, "Objective bayesian approach for binary hypothesis testing of multivariate gaussian observations," *IEEE Transactions on Information Theory*, vol. 69, no. 2, pp. 1337–1354, Feb. 2023.

[4] V. W. Neo, S. Weiss, S. W. McKnight, A. O. T. Hogg, and P. A. Naylor, "Polynomial eigenvalue decomposition-based target speaker voice activity detection in the presence of competing talkers," in *Int. Workshop on Acoustic Signal Enhancement*, Bamberg, Germany, Sep. 2022.

[5] V. W. Neo, S. Weiss, and P. A. Naylor, "A polynomial subspace projection approach for the detection of weak voice activity," in *Sensor Signal Processing for Defence*, London, UK, Sep. 2022, pp. 1–5.

[6] S. Weiss, C. Delaosa, J. Matthews, I. Proudler, and B. Jackson, "Detection of weak transient signals using a broadband subspace approach," in *Sensor Signal Processing for Defence*, Edinburgh, Scotland, Sep. 2021, pp. 65–69.

[7] B. Friedlander and B. Porat, "Detection of transient signals by the gabor representation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 2, pp. 169–180, Feb. 1989.

[8] ——, "Performance analysis of transient detectors based on a class of linear data transforms," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 665–673, Mar. 1992.

[9] B. Porat and B. Friedlander, "Performance analysis of a class of transient detection algorithms-a unified framework," *IEEE Transactions on Signal Processing*, vol. 40, no. 10, pp. 2536–2546, Oct. 1992.

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

[11] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Processing*, vol. 42, no. 8, pp. 2146–2157, Aug. 1994.

[12] P. Strobach, "Low rank detection of multichannel gaussian signals using a constrained inverse," in *IEEE Int. Conference on Acoustics, Speech and Signal Processing*, vol. iv, Apr. 1994, pp. 245–248.

[13] D. Lundstrom, M. Viberg, and A. M. Zoubir, "Multiple transient estimation using bootstrap and subspace methods," in *9th IEEE Signal Processing Workshop on Statistical Signal and Array Processing*, Sep. 1998, pp. 184–187.

[14] Zhen Wang and P. Willett, "A performance study of some transient detectors," *IEEE Transactions on Signal Processing*, vol. 48, no. 9, pp. 2682–2685, Sep. 2000.

[15] Zhen Wang and P. K. Willett, "All-purpose and plug-in power-law detectors for transient signals," *IEEE Transactions on Signal Processing*, vol. 49, no. 11, pp. 2454–2466, Nov. 2001.

[16] F. Labeau, J.-C. Chiang, M. Kieffer, P. Duhamel, L. Vandendorpe, and B. Macq, "Oversampled filter banks as error correcting codes: theory and impulse noise correction," *IEEE Transactions on Signal Processing*, vol. 53, no. 12, pp. 4619–4630, 2005.

[17] O. Besson, "Maximum likelihood covariance matrix estimation from two possibly mismatched data sets," *Signal Processing*, vol. 167, p. 107285, 2020.

[18] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*, 4th ed. Springer, 2022.

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

[20] S. Weiss, J. Pestana, and I. K. Proudler, "On the existence and uniqueness of the eigenvalue decomposition of a parahermitian matrix," *IEEE Trans. Signal Processing*, vol. 66, no. 10, pp. 2659–2672, May 2018.

[21] S. Weiss, J. Pestana, I.K. Proudler, and F.K. Coutts, "Corrections to "on the existence and uniqueness of the eigenvalue decomposition of a parahermitian matrix"," *IEEE Transactions on Signal Processing*, vol. 66, no. 23, pp. 6325–6327, Dec. 2018.

[22] G. Barbarino and V. Noferini, "On the Rellich eigendecomposition of para-Hermitian matrices and the sign characteristics of $*$-palindromic matrix polynomials," *Linear Algebra Appl.*, vol. 672, pp. 1–27, Sep. 2023.

[23] S. Weiss, I. Proudler, F. Coutts, and F. Khattak, "Eigenvalue decomposition of a parahermitian matrix: extraction of analytic eigenvectors," *IEEE Transactions on Signal Processing*, vol. 71, pp. 1642–1656, Apr. 2023.

[24] V. Neo, S. Redif, J. McWhirter, J. Pestana, I. Proudler, S. Weiss, and P. Naylor, "Polynomial eigenvalue decomposition for multichannel broadband signal processing," *IEEE Signal Processing Magazine*, vol. 40, no. 7, pp. 18–37, Nov. 2023.

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

[26] 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.

[27] C. Delaosa, J. Pestana, N. J. Goddard, S. Somasundaram, and S. Weiss, "Sample space-time covariance matrix estimation," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Brighton, UK, May 2019, pp. 8033–8037.