SAMPLE SPACE-TIME COVARIANCE MATRIX ESTIMATION

Connor Delaosa[1], *Jennifer Pestana*[2], *Nicholas J. Goddard*[3], *Samuel Somasundaram*[4], *Stephan Weiss*[1]

[1] Dept. of Electronic & Electrical Eng., University of Strathclyde, Glasgow, Scotland
[2] Dept. of Mathematics & Statistics, University of Strathclyde, Glasgow, Scotland
[3] DSTL Portsdown, Fareham, Hampshire, UK
[4] Thales Underwater Systems, Cheadle, Stockport, UK

ABSTRACT

Estimation errors are incurred when calculating the sample space-time covariance matrix. We formulate the variance of this estimator when operating on a finite sample set, compare it to known results, and demonstrate its precision in simulations. The variance of the estimation links directly to previously explored perturbation of the analytic eigenvalues and eigenspaces of a parahermitian cross-spectral density matrix when estimated from finite data.

***Index Terms*—** Space-time covariance, estimation, parahermitian matrix EVD, polynomial matrices.

1. INTRODUCTION

A parahermitian matrix — typically a cross-spectral density (CSD) matrix emerging as the z-transform of a space-time covariance matrix — can be decomposed into a product of analytic paraunitary matrices and a diagonalised parahermitian matrix [1] with few exceptions [2]. A spectrally majorised, not necessarily analytic version of this factorisation is the McWhirter decomposition [3], which approximates the factorisation by polynomial paraunitary and diagonal parahermitian matrices. A number of algorithms for the latter have emerged [3–10] and in turn triggered various applications ranging from broadband multiple-input and multiple-output (MIMO) systems [11, 12], to coding [13], beamforming [14, 15], source separation [16] and angle of arrival estimation [17, 18], to name but a few.

In applications, the space-time covariance or the CSD matrix generally have to be estimated from data. While the accuracy of the decomposition itself has been investigated in [19, 20], and limiting factors due to algorithm-internal order reductions [8, 21–23] and the conditioning of the underlying source model [24] are known, it is only recently that the effect of estimating the space time covariance matrix from a finite data set has been addressed [25]. While [25] linked the estimation error to the perturbation of the eigenvectors and eigenspaces of the CSD matrix, the formulation of this estimation error had still been missing.

This paper aims to close the gap, in order to link eigenvalue and -space perturbations directly to the variance of sample space-time covariance. To date, results have been derived for the broadband single channel case, i.e. for the sample auto-correlation sequence. Various attempts have been undertaken for random signals that can be modelled as first order auto-regressive processes [26, 27], or generally [28, 29]. For the broadband case, analysis has generally been restricted to narrowband signals, such that the spatial covariance matrix is Wishart distributed [30, 31]. We derive the variance of a sample cross-correlation sequence, which then forms the building block for the sample space-time covariance. Particularisation of our results agree with [28, 30, 31] and results from spectral estimation such as [32].

We commence with a definition of the space-time covariance matrix, and review its properties and matrix factorisation in Sec. 2. In Sec. 3 we analyse the sample cross-correlation sequence, which expands to the space-time covariance in Sec. 4, followed by experimental verification in Sec. 5.

2. SPACE-TIME COVARIANCE MATRIX AND PARAHERMITIAN MATRIX EVD

2.1. Space-Time Covariance

Given M sensor measurements $x_m[n]$, $m = 1 \dots M$, organised in a vector $\mathbf{x}[n] = [x_1[n] \dots x_m[n]]^{\mathrm{T}}$, the space-time covariance matrix of the data is defined as $\mathbf{R}[\tau] = \mathcal{E}\{\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n-\tau]\}$, with $\mathcal{E}\{\cdot\}$ the expectation operator. The source model or innovation filter [33] in Fig. 1 ties this data vector $\mathbf{x}[n]$ to L zero-mean unit-variance mutually independent complex circularly symmetric Gaussian sources u_ℓ, $\ell = 1 \dots L$, such that $\mathcal{E}\{u_\ell[n]u_\nu[n-\tau]\} = \delta[\tau]\delta[\ell-\nu]$ for $\nu = 1 \dots L$ [34]. As a result, the space-time covariance matrix can be expressed as

$$\mathbf{R}[\tau] = \sum_n \mathbf{H}[n]\mathbf{H}^{\mathrm{H}}[n-\tau] \ ,$$

This work was supported in parts by the Engineering and Physical Sciences Research Council (EPSRC) Grant number EP/S000631/1 and the MOD University Defence Research Collaboration in Signal Processing, and a John Anderson Research Award by the University of Strathclyde.

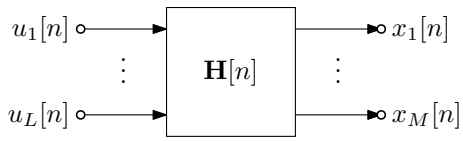

Fig. 1. Source model for M convolutively mixed signals arising from L independent unit-variance zero-mean sources.

where $\mathbf{H}[n] \in \mathbb{C}^{M\times L}$ is a matrix of filters. If the entry in the mth row and ℓth column of $\mathbf{H}[n]$ represents the impulse response $h_{m\ell}[n]$, then

$$r_{m\mu}[\tau] = \sum_n \sum_{\ell=1}^{L} h_{m\ell}[n] h^*_{\mu\ell}[n-\tau] \tag{1}$$

is a cross-correlation sequence that occupies the mth row and μth column of $\mathbf{R}[\tau]$, with $\{\cdot\}^*$ denoting complex conjugation.

2.2. Cross-Spectral Density

Since the space-time covariance matrix comprises auto- and cross-correlation sequences, it satisfies the symmetry $\mathbf{R}[\tau] = \mathbf{R}^{\mathrm{H}}[-\tau]$. Its z-transform, the cross-spectral density (CSD) matrix $\boldsymbol{R}(z) = \sum_\tau \mathbf{R}[\tau] z^{-\tau}$ — or short $\boldsymbol{R}(z) \bullet\!\!-\!\!\circ\, \mathbf{R}[\tau]$ to denote a transfom pair — therefore is a parahermitian matrix, such that its parahermitian transpose, denoted by the operator $\{\cdot\}^{\mathrm{P}}$, is equal to itself: $\boldsymbol{R}^{\mathrm{P}}(z) = \boldsymbol{R}^{\mathrm{H}}(1/z^*) = \boldsymbol{R}(z)$ [35].

2.3. Parahermitian Matrix EVD

A parahermitian, analytic $\mathbf{R}(z)$ admits a parahermitian matrix EVD (PhEVD) [1]

$$\mathbf{R}(z) = \mathbf{U}(z)\boldsymbol{\Lambda}(z)\mathbf{U}^{\mathrm{P}}(z)\ , \tag{2}$$

where $\mathbf{U}(z)$ is a paraunitary matrix of eigenvectors and $\boldsymbol{\Lambda}(z)$ is a diagonal parahermitian matrix of eigenvalues. In most standard cases, these factors can be selected to be analytic [2]. This is closely related to the McWhirter decomposition [3], where the factors $\boldsymbol{U}(z)$ and a spectrally majorised $\boldsymbol{\Lambda}(z)$ are approximated by polynomials, i.e. are of finite order, while the terms on the r.h.s. of (2) are generally algebraic or transcendental.

If the space time covariance matrix is estimated from a finite set of samples, the obtained matrix $\hat{\mathbf{R}}[\tau] \circ\!\!-\!\!\bullet\, \hat{\boldsymbol{R}}(z)$ will differ from $\boldsymbol{R}(z)$. Similarly, the PhEVD $\hat{\mathbf{R}}(z) = \hat{\mathbf{U}}(z)\hat{\boldsymbol{\Lambda}}(z)\hat{\mathbf{U}}^{\mathrm{P}}(z)$ will deviate from that in (2). In [25], the norm of the modelling error

$$\boldsymbol{E}(z) = \boldsymbol{R}(z) - \hat{\boldsymbol{R}}(z) \tag{3}$$

was linked to the deviation in the eigenvalues, i.e. the difference between $\boldsymbol{\Lambda}(z)$ and $\hat{\boldsymbol{\Lambda}}(z)$. The perturbation of the subspace angle between a particular eigenspace of $\boldsymbol{U}(z)$ and of $\hat{\boldsymbol{U}}(z)$ can in turn be linked to the norm of $\boldsymbol{E}(z)$ and the distance to the nearest eigenvalue, i.e. near eigenvalues with algebraic multiplicity greater than one, subspaces can undergo a larger perturbation [25].

3. SAMPLE CROSS-CORRELATION SEQUENCE

In applications, the space-time covariance matrix must be estimated from data. If only a set of N snap-shots of $\mathbf{x}[n]$, $n = 0\ldots(N-1)$, is available, then generally the estimate for the space-time covariance matrix, $\hat{\mathbf{R}}[\tau]$, will be prone to estimation errors. Since the cross-correlation sequence in (1) is the most general entry of the space-time covariance matrix, we focus on its estimation in order to explore the estimation of $\mathbf{R}[\tau]$.

3.1. Unbiased Estimator

The cross-correlation sequence between two signals $x_m[n]$ and $x_\mu[n]$, $m, \mu \in \{1\ldots M\}$, is defined as

$$r_{m\mu}[\tau] = \mathcal{E}\big\{x_m[n]x^*_\mu[n-\tau]\big\}\ . \tag{4}$$

Assuming ergodicity and therefore by implication stationarity for the involved signals, the estimation of $r_{m\mu}[\tau]$ over a set of N time snapshots,

$$\hat{r}_{m\mu}[\tau] = \begin{cases} \frac{1}{N-|\tau|} \sum\limits_{n=0}^{N-|\tau|-1} x_m[n+\tau]x^*_\mu[n]\ , & \tau \geq 0 \\ \frac{1}{N-|\tau|} \sum\limits_{n=0}^{N-|\tau|-1} x_m[n]x^*_\mu[n-\tau]\ , & \tau < 0 \end{cases} \tag{5}$$

can be shown to be unbiased. For example for $\tau \geq 0$,

$$\begin{aligned} \text{mean}\{\hat{r}_{m\mu}[\tau]\} &= \mathcal{E}\{\hat{r}_{m\mu}[\tau]\} \\ &= \frac{1}{N-|\tau|} \sum_{n=0}^{N-\tau-1} \mathcal{E}\big\{x_m[n]x^*_\mu[n-\tau]\big\} \\ &= \frac{1}{N-|\tau|} \sum_{n=0}^{N-\tau-1} r_{m\mu}[\tau] = r_{m\mu}[\tau]\ , \end{aligned}$$

i.e. the quantity estimated via (5) tends towards the true cross-correlation sequence defined in (4).

3.2. Variance

The variance of the cross-correlation sequence estimator is given by

$$\begin{aligned} \text{var}\{\hat{r}_{m\mu}[\tau]\} &= \mathcal{E}\{(\hat{r}_{m\mu}[\tau] - r_{m\mu}[\tau])(\hat{r}_{m\mu}[\tau] - r_{m\mu}[\tau])^*\} \\ &= \mathcal{E}\big\{\hat{r}_{m\mu}[\tau]\hat{r}^*_{m\mu}[\tau]\big\} - \mathcal{E}\{\hat{r}_{m\mu}[\tau]\}\, r^*_{m\mu}[\tau] - \\ &\quad - r_{m\mu}[\tau]\mathcal{E}\big\{\hat{r}^*_{m\mu}[\tau]\big\} + r_{m\mu}[\tau]r^*_{m\mu}[\tau] \\ &= \mathcal{E}\big\{\hat{r}_{m\mu}[\tau]\hat{r}^*_{m\mu}[\tau]\big\} - r_{m\mu}[\tau]r^*_{m\mu}[\tau]\ . \end{aligned} \tag{6}$$

Inserting the estimation in (5) into (6), we obtain fourth order terms. For Gaussian signals, the cumulants of order three and above are zero [36, 37], which also holds for the complex-valued case [38], such that

$$\begin{aligned}\mathcal{E}\{x_m[n]x_\mu^*[n-\tau]x_m^*[n]x_\mu[n-\tau]\} = \\ \mathcal{E}\{x_m[n]x_\mu^*[n-\tau]\}\cdot\mathcal{E}\{x_m^*[n]x_\mu[n-\tau]\} \\ +\mathcal{E}\{x_m[n]x_m^*[n]\}\cdot\mathcal{E}\{x_\mu^*[n-\tau]x_\mu[n-\tau]\} \\ +\mathcal{E}\{x_m[n]x_\mu[n-\tau]\}\cdot\mathcal{E}\{x_\mu^*[n-\tau]x_m^*[n]\}\ .\end{aligned}$$

Therefore, for $\tau \geq 0$, the variance of the estimator in (5) becomes

$$\begin{aligned}\mathrm{var}\{\hat{r}_{m\mu}[\tau]\} &= \frac{1}{(N-|\tau|)^2}\sum_{n,\nu=0}^{N-|\tau|-1}\big(\mathcal{E}\{x_m[n+\tau]x_\mu^*[n]\}\cdot \\ &\qquad\cdot\mathcal{E}\{x_m^*[\nu+\tau]x_\mu[\nu]\}+ \\ &\quad+\mathcal{E}\{x_m[n+\tau]x_m^*[\nu+\tau]\}\,\mathcal{E}\{x_\mu^*[n]x_\mu[\nu]\} \\ &\quad+\mathcal{E}\{x_m[n+\tau]x_\mu[\nu]\}\,\mathcal{E}\{x_\mu^*[n]x_\mu^*[\nu+\tau]\}\big) \\ &\qquad-r_{m\mu}[\tau]r_{m\mu}^*[\tau] \\ &= \frac{1}{(N-|\tau|)^2}\sum_{n,\nu=0}^{N-|\tau|-1}\big(\mathcal{E}\{x_m[n]x_m^*[\nu]\}\cdot \\ &\qquad\cdot\mathcal{E}\{x_\mu^*[n]x_\mu[\nu]\}+ \\ &\quad+\mathcal{E}\{x_m[n]x_\mu[\nu-\tau]\}\,\mathcal{E}\{x_m^*[\nu]x_\mu^*[n-\tau]\}\big)\ .\end{aligned} \tag{7}$$

The same result can be obtained for $\tau < 0$, and matches results reached in [32].

Note that the first term in (7) can be simplified as

$$\begin{aligned}&\sum_{n,\nu=0}^{N-|\tau|-1}\big(\mathcal{E}\{x_m[n]x_m^*[\nu]\}\,\mathcal{E}\{x_\mu^*[n]x_\mu[\nu]\}\big) \\ &= \sum_{n,\nu=0}^{N-|\tau|-1}\big(\mathcal{E}\{x_m[n]x_m^*[n-(n-\nu)]\}\cdot \\ &\qquad\qquad\cdot\mathcal{E}\{x_\mu^*[n]x_\mu[n-(n-\nu)]\}\big) \\ &= \sum_{n,\nu=0}^{N-|\tau|-1} r_{mm}[n-\nu]r_{\mu\mu}^*[n-\nu] \\ &= \sum_{t=-N+|\tau|+1}^{N-|\tau|-1}(N-|\tau|-|t|)r_{mm}[t]r_{\nu\nu}^*[t]\ .\end{aligned}$$

A similar method works for the second term in (7). With $\bar{r}_{m\mu}[\tau] = \mathcal{E}\{x_m[n]x_\mu[n-\tau]\}$ denoting the complementary cross-correlation sequence, the variance of the sample cross-correlation sequence becomes

$$\begin{aligned}\mathrm{var}\{\hat{r}_{m\mu}[\tau]\} = \frac{1}{(N-|\tau|)^2}\sum_{t=-N+|\tau|+1}^{N-|\tau|-1}(N-|\tau|-|t|)\cdot \\ \cdot\big(r_{mm}[t]r_{\mu\mu}^*[t]+\bar{r}_{m\mu}[\tau+t]\bar{r}_{m\mu}^*[\tau-t]\big)\ .\end{aligned} \tag{8}$$

If $u[n]$ is complex valued with a circularly symmetric distribution, then given the source model in Fig. 1, $\bar{r}_{xy}[\tau] = 0\ \forall\tau\in\mathbb{Z}$. Nevertheless, we continue to carry the term in order to particularise the result in (8) to the real valued case.

3.3. Particularisation

The result in (8) generalises a number of solutions reported in the literature. If $\mathbf{u}[n]\in\mathbb{R}^L$ and $m=\mu$, then (8) simplifies to

$$\begin{aligned}\mathrm{var}\{\hat{r}_{mm}[\tau]\} = \frac{1}{(N-|\tau|)^2}\sum_{t=-N+|\tau|+1}^{N-|\tau|-1}(N-|\tau|-|t|)\cdot \\ \cdot\big(|r_{mm}[t]|^2+r_{mm}[\tau+t]r_{mm}[\tau-t]\big)\ .\end{aligned}$$

This matches with the result reported in [28].

If the transfer function $\boldsymbol{H}(z)\circ\!\!-\!\!\bullet\,\mathbf{H}[n]$ is a constant matrix, $\boldsymbol{H}(z)=\mathbf{H}_0$, then the signals $x_m[n]$ and $x_\mu[n]$ only have non-zero correlation for the instant case $\tau=0$. If further $\mathbf{u}[n]\in\mathbb{R}^L$ and $\mathbf{H}_0\in\mathbb{R}^{M\times L}$, then the space-time covariance $\mathbf{R}[\tau]=\mathbf{H}_0\mathbf{H}_0^{\mathrm{T}}\delta[\tau]$ is Wishart-distributed. For the instantaneous and real case, (8) simplifies to

$$\mathrm{var}\{\hat{r}_{m\mu}[0]\} = \frac{1}{N}\left(r_{mm}[0]r_{\mu\mu}[0]+|r_{m\mu}[0]|^2\right)\ ,$$

which indeed matches the variance of a Wishart distribution.

4. SAMPLE SPACE-TIME COVARIANCE

4.1. Sample Space-Time Covariance Error

Assume that the space-time covariance matrix has support of length $2\tau_{\max}+1$, i.e. $\mathbf{R}[\tau]=\mathbf{0}\ \forall|\tau|>\tau_{\max}$. Further assume that $\hat{\mathbf{R}}[\tau]$ is estimated over a support length of $2T+1$. Abbreviating $\mathbf{E}[\tau]=\mathbf{R}[\tau]-\hat{\mathbf{R}}[\tau]$, the mean square error is

$$\begin{aligned}\xi &= \mathcal{E}\left\{\sum_{\tau=-\infty}^{\infty}\|\mathbf{E}[\tau]\|_{\mathrm{F}}^2\right\} \\ &= \underbrace{\sum_{\tau=-T}^{T}\mathcal{E}\{\|\mathbf{E}[\tau]\|_{\mathrm{F}}^2\}}_{\xi_1}\ +\ \underbrace{2\sum_{\tau=T+1}^{\tau_{\max}}\|\mathbf{R}[\tau]\|_{\mathrm{F}}^2}_{\xi_2}\ ,\end{aligned}$$

where the first term, ξ_1, is an estimation error due to (8), while the second term, ξ_2, represents a truncation error. Note that $\xi_2=0$ if $T\geq\tau_{\max}$.

For the estimation error, using (8),

$$\begin{aligned}\xi_1 &= \sum_{\tau=-T}^{T}\sum_{m,\mu=1}^{M}\mathrm{var}\{\hat{r}_{m\mu}[\tau]\} \\ &= \sum_{\tau=-T}^{T}\sum_{t=-N+|\tau|+1}^{N-|\tau|-1}\frac{N-|\tau|-|t|}{(N-|\tau|)^2}\left(|\mathrm{tr}\{\mathbf{R}[t]\}|^2+\right. \\ &\qquad\left.+\ \mathrm{vec}\{\mathbf{R}[\tau-t]\}^{\mathrm{H}}\mathrm{vec}\{\mathbf{R}[\tau+t]\}\right)\ ,\end{aligned} \tag{9}$$

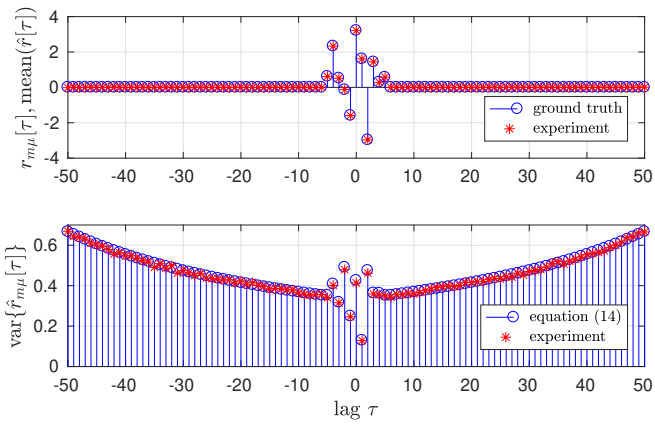

Fig. 2. (top) ground truth and mean sample cross-correlation sequence, and (bottom) its variance, both calculated according to (8) and estimated from real valued data.

where the operator $\mathrm{vec}\{\cdot\}$ vectorises its argument and T is the support of the estimate. Therefore, the modelling error ξ depends only on the space-time covariance matrix itself, the sample size N, and the support of the estimate, T.

4.2. Optimum Support Length

Note that (9) also presents a formulation for the expected mean square value of the error $\boldsymbol{E}(z) \bullet\!\!-\!\!\circ\, \mathbf{E}[\tau]$ in (3), which forms the basis for the perturbation analysis of the eigenvalues and eigenvector in [25]. Since the sample size N and the specific ground truth $\mathbf{R}[\tau]$ are given for a particular problem, the only way to minimise this perturbation is the judicious selection of the range of lags, $|\tau| \leq T$, over which $\hat{\mathbf{R}}[\tau]$ is evaluated. The optimum value T_{opt} for T in terms of the minimum perturbation is therefore

$$T_{\mathrm{opt}} = \arg\min_{T} \xi \ .$$

In general, this will be a trade-off between the terms ξ_1 and ξ_2. Since $T > \tau_{\max}$ leads to $\xi_2 = 0$, and ξ_1 generally grows with increasing T, we find $T_{\mathrm{opt}} < \tau_{\max}$, i.e. it appears better to underestimate than to overestimate the support of $\mathbf{R}[\tau]$ in practise.

5. SIMULATIONS AND RESULTS

We first demonstrate the accuracy of (8) for the variance of a sample cross-correlation sequence. For an arbitrary given cross-correlation by means of an innovation filter model of order 5 with a single source, $L = 1$ in Fig. 1 and (1), for $N = 100$ the theoretical (8) is compared to the mean variance over an ensemble of size 10^4 in Figs. 2 and 3 for a real- and complex-valued scenarios.

To check the accuracy of the expected estimation error (9), an $\mathbf{R}[\tau]$ of order 100 is generated by the source model in [7]. Fig. 4 compares results over an ensemble of 10^4 sample sets, each $L = 500$ long, to the theoretical values. These match well, and also demonstrate that in this case $T_{\mathrm{opt}} = 10$ is significantly shorter than $\tau_{\max} = 50$.

6. CONCLUSIONS

This paper has investigated the dependencies of the modelling error that is incurred when estimating a space-time covariance matrix from a finite sample set — this is affected by the size of the set, but also the ground truth space-time covariance matrix. The derived expressions match what has previously been identified for sample auto-correlation sequences for the case of temporal correlation only, and to the Wishart distribution in the case of spatial correlation only; in simulations, we have also demonstrated a close match to experiments.

The mean square modelling error is a metric that has previously been established in [25] to perturb the eigenvalues and eigenspaces of the space-time covariance matrix; therefore, the results, particularly (9), now directly link this perturbation to the sample size and the ground truth matrix.

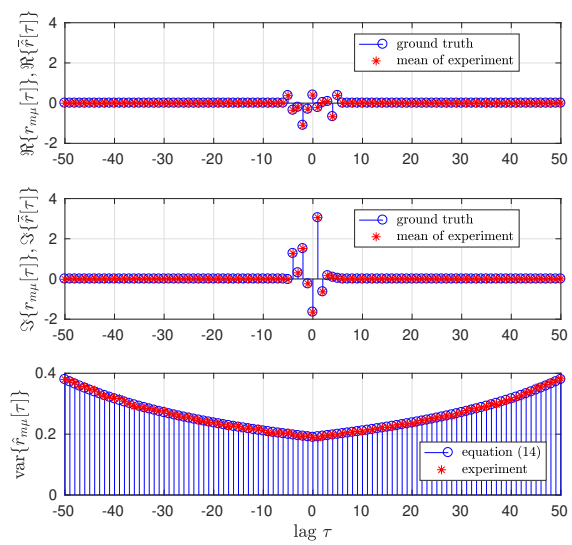

Fig. 3. Complex valued equivalent to Fig. 2, with (top) real part, (middle) imaginary part, and (bottom) variance.

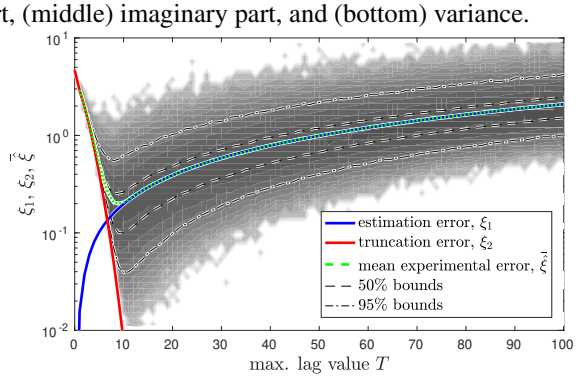

Fig. 4. Comparison of the ensemble modelling error to truncation and expected estimation errors.

7. REFERENCES

[1] 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.

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

[3] 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.

[4] A. Tkacenko, "Approximate eigenvalue decomposition of para-hermitian systems through successive FIR paraunitary transformations," in *IEEE ICASSP*, Dallas, TX, Mar. 2010, pp. 4074–4077.

[5] S. Redif, J.G. 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.

[6] M. Tohidian, H. Amindavar, and A.M. Reza, "A DFT-based approximate eigenvalue and singular value decomposition of polynomial matrices," *EURASIP J. Adv. Signal Processing*, **2013**(1), pp. 1–16, 2013.

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

[8] J. Corr, K. Thompson, S. Weiss, I.K. Proudler, and J.G. McWhirter, "Row-shift corrected truncation of paraunitary matrices for PEVD algorithms," in *EUSIPCO*, Nice, France, Aug. 2015, pp. 849–853.

[9] F.K. Coutts, K. Thompson, S. Weiss, and I.K. Proudler, "A comparison of iterative and DFT-based polynomial matrix eigenvalue decompositions," in *IEEE CAMSAP*, Curacao, Dec. 2017.

[10] 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.

[11] C.H. Ta and S. Weiss, "A jointly optimal precoder and block decision feedback equaliser design with low redundancy," in *EUSIPCO*, Poznan, Poland, Sep. 2007, pp. 489–492.

[12] S. Weiss, N. J. Goddard, S. Somasundaram, I. K. Proudler, and P. A. Naylor, "Identification of broadband source-array responses from sensor second order statistics," in *SSPD*, London, UK, Dec. 2017, pp. 1–5.

[13] S. Weiss, S. Redif, T. Cooper, C. Liu, P. Baxter, and J.G. McWhirter, "Paraunitary oversampled filter bank design for channel coding," *EURASIP J. Adv. SP*, **2006**:1–10, 2006.

[14] S. Redif, J.G. McWhirter, P.D. Baxter, and T. Cooper, "Robust broadband adaptive beamforming via polynomial eigenvalues," in *OCEANS*, Boston, MA, Sep. 2006.

[15] S. Weiss, S. Bendoukha, A. Alzin, F.K. Coutts, I.K. Proudler, and J.A. Chambers, "MVDR broadband beamforming using polynomial matrix techniques," in *EUSIPCO*, Nice, France, Sep. 2015, pp. 839–843.

[16] S. Redif, S. Weiss, and J.G. McWhirter, "Relevance of polynomial matrix decompositions to broadband blind signal separation," *Signal Processing*, **134**:76–86, May 2017.

[17] M. Alrmah, S. Weiss, and S. Lambotharan, "An extension of the MUSIC algorithm to broadband scenarios using polynomial eigenvalue decomposition," in *EUSIPCO*, Barcelona, Spain, Aug. 2011, pp. 629–633.

[18] S. Weiss, M. Alrmah, S. Lambotharan, J.G. McWhirter, and M. Kaveh, "Broadband angle of arrival estimation methods in a polynomial matrix decomposition framework," in *IEEE CAMSAP*, Dec. 2013, pp. 109–112.

[19] M.A. Alrmah, J. Corr, A. Alzin, K. Thompson, and S. Weiss, "Polynomial subspace decomposition for broadband angle of arrival estimation," in *SSPD*, Edinburgh, Scotland, Sep. 2014.

[20] F.K. Coutts, K. Thompson, S. Weiss, and I.K. Proudler, "Impact of fast-converging PEVD algorithms on broadband AoA estimation," in *SSPD*, London, UK, Dec. 2017.

[21] J. Foster, J.G. McWhirter, and J.A. Chambers, "Limiting the order of polynomial matrices within the SBR2 algorithm," in *IMA Maths in Sig. Proc.*, Cirencester, UK, Dec. 2006

[22] C.H. Ta and S. Weiss, "Shortening the order of paraunitary matrices in SBR2 algorithm," in *ICICSP*, Singapore, Dec. 2007.

[23] J. Corr, K. Thompson, S. Weiss, I.K. Proudler, and J.G. McWhirter, "Shortening of paraunitary matrices obtained by polynomial eigenvalue decomposition algorithms," in *SSPD*, Edinburgh, Scotland, Sep. 2015.

[24] J. Corr, K. Thompson, S. Weiss, I.K. Proudler, and J.G. McWhirter, "Impact of source model matrix conditioning on PEVD algorithms," in *IET/EURASIP ISP*, London, UK, Dec. 2015.

[25] C. Delaosa, F.K. Coutts, J. Pestana, and S. Weiss, "Impact of space-time covariance estimation errors on a parahermitian matrix EVD," in *IEEE SAM*, Sheffield, UK, July 2018.

[26] R.L. Anderson, "Distribution of the serial correlation coefficient," *Ann. Math. Statist.*, **13**(1):1–13, 1942.

[27] H. Goldsmith, "The exact distributions of the serial correlation coefficients and an evaluation on some approximate distributions," *J. Stat. Comp. and Sim.*, **5**(2,):115–134, 1977.

[28] S. Kay, "The effect of sampling rate on autocorrelation estimation," *IEEE Trans. ASSP*, **29**(4):859–867, Aug. 1981.

[29] A. Steinhardt and J. Makhoul, "On the autocorrelation of finite-length sequences," *IEEE Trans. Acoustics, Speech, and Signal Processing*, bf 33(6):1516–1520, Dec. 1985.

[30] R.K. Mallik, "The pseudo-Wishart distribution and its application to MIMO systems," *IEEE Trans. IT*, bf 49(10):2761–2769, Oct. 2003.

[31] L. Zhuang and A.T. Walden, "Sample mean versus sample Fréchet mean for combining complex Wishart matrices: A statistical study," *IEEE Trans. SP*, **65**(17):4551–4561, Sep. 2017.

[32] D.D. Ariananda and G. Leus, "Compressive wideband power spectrum estimation," *IEEE Trans. SP*, **60**(9):4775–4789, Sep. 2012.

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

[34] B. Picinbono, "On circularity," *IEEE Trans. Signal Processing*, **42**(12):3473–3482, Dec. 1994.

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

[36] W. Bär and F. Dittrich, "Useful formula for moment computation of normal random variables with nonzero means," *IEEE Trans. AC*, **16**(3):263–265, June 1971.

[37] J.M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *Proc. IEEE*, **79**(3):278–305, Mar. 1991.

[38] P.J. Schreier and L.L. Scharf, *Statistical Signal Processing of Complex-Valued Data*, Cambridge Univ. Press, 2010.