Knowledge-Aided Covariance Matrix Estimation via Kronecker Product Expansions for Airborne STAP

Guohao Sun, *Student Member, IEEE*, Zishu He, *Member, IEEE*, Jun Tong, and Xuejing Zhang, *Student Member, IEEE*

***Abstract*— This letter proposes a new approach for knowledge-aided estimation of structured clutter covariance matrices (CCMs) in airborne radar systems with limited training data. First, we model the CCM in space–time adaptive processing (STAP) as a sum of low-rank Kronecker products. We then apply a permutation operation to convert the Kronecker factors into linear structures and propose a novel CCM estimation method under the maximum-likelihood framework. Employing a proximal gradient algorithm, the proposed method simultaneously exploits the knowledge about the clutter and the Kronecker structure of the CCM. We finally evaluate the performance of the proposed method using real data from airborne STAP.**

***Index Terms*— Covariance matrix estimation, knowledge aided (KA), space–time adaptive processing (STAP), sum of low-rank Kronecker products.**

I. Introduction

SPACE–time adaptive processing (STAP) is an effective tool for detecting slowly varying targets, especially when ground clutter with a wide Doppler spectrum is present [1]–[3]. Accurate estimation of the clutter covariance matrix (CCM) using training data from range gates close to the cell under test (CUT) is crucial for STAP. The accuracy of the CCM estimation depends on the number of homogenous training measurements, which can be small in practice, especially in urban, mountain, and other complex scenarios [3], [4]. On the other hand, the dimensionality of the CCM in airborne radar systems can be very large due to the usage of large antenna arrays and coherent pulse trains [5], [6]. This leads to the challenge of small sample supports for high-dimensional CCM estimation. Signal processing and statistical learning techniques are exploited to enhance the estimation of CCM, among which knowledge-aided (KA) and structured CCM estimators have proven to be effective solutions.

Exploiting *a priori* information, KA methods can improve the performance of CCM estimation and adaptive detection for synthetic aperture radar imaging, terrain data, digital elevation, etc. [7]. The Bayesian approach is advocated as an effective method for KA CCM estimation [8], [9], exploiting a joint distribution of the CCM and the KA covariance matrix. In particular, the maximum *a priori* estimate of the CCM for the CUT is derived in [9]. It is shown that the estimator amounts to KA colored loading (KACL) of the sample covariance matrix (SCM) of the training data [10], [11].

Exploiting the structure of CCM, e.g., the low rankness, leads to a reduced number of unknown parameters and provides another approach to address the challenge of insufficient training data [12]–[15]. In [15], the CCM is modeled using the Kronecker product of two low-dimensional low-rank covariance matrices, and maximum-likelihood (ML) estimation exploiting this structure has been proposed. In airborne radar systems, the CCM can be modeled as a sum of low-rank Kronecker products [16], [17], which has been confirmed by the principal component analysis (PCA) [16], [18]. A permuted singular value thresholding (PSVT) algorithm has been proposed to exploit this structure, which leads to a faster convergence rate than the standard SCM [19].

This letter considers the KA ML estimation of CCM under the model of Kronecker product expansions. A permutation operation is adopted to tackle the structure constraint of CCM. The proposed method has the following features.

1) Adopting the permutation operation turns a sum of Kronecker products into a sum of vector outer products.
2) The permutation is invertible. For a matrix that has the structure of vector outer product expansions, the inverse permutation follows a Kronecker product expansion.
3) The number of large eigenvalues of the permuted CCM, which account for most of the energy, is smaller than the original CCM. Therefore, the permutation reduces the number of unknown parameters to be estimated.

In addition, a relaxation of the ML problem is adopted, which results in a convex optimization problem [20] to be solved by a proximal gradient algorithm. The proposed estimator is evaluated for the STAP application in airborne radar using real data, which shows notable improvement over traditional estimators.

The rest of this letter is organized as follows. Section II describes the structured covariance matrix and formulates the CCM estimation problem. Section III introduces the proposed estimator for STAP. Experiments and analysis of the proposed estimator are given in Section IV. Conclusions are drawn in Section V.

Manuscript received January 18, 2018; revised January 22, 2018; accepted January 25, 2018. Date of publication February 12, 2018; date of current version March 23, 2018. This work was supported by the National Natural Science Foundation of China under Grant 61671139. *(Corresponding author: Guohao Sun.)*

G. Sun, Z. He, and X. Zhang are with the School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: sghsjw2015@gmail.com).

J. Tong is with the School of Electrical, Computer, and Telecommunications Engineering, University of Wollongong, Wollongong, NSW 2522, Australia.

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

Digital Object Identifier 10.1109/LGRS.2018.2799329

II. Problem Statement

A. Structure of Airborne Radar CCM

Consider an airborne radar system that employs an N-element uniform linear array and M-pulse trains in a coherent processing interval. Assuming that there are N_c independent clutter patches, the received CCM can be described as

$$\mathbf{R}_c = \sum_{i=1}^{N_c} \xi_i \left(\mathbf{\Phi}_i \circ \mathbf{a}_i \mathbf{a}_i^\dagger\right) \otimes \left(\mathbf{\Gamma}_i \circ \mathbf{b}_i \mathbf{b}_i^\dagger\right) \quad (1)$$

where $\circ$ and $\otimes$ denote the Hadamard and Kronecker products, respectively, $\dagger$ is the conjugate transpose, ξ_i is the power of the ith clutter patch, $\mathbf{a}_i$ and $\mathbf{b}_i$ are the spatial and temporal steering vectors of the ith clutter patch, respectively, $\mathbf{\Gamma}_i$ is the taper for the ith clutter patch covariance matrix accounting for internal clutter motion, and $\mathbf{\Phi}_i$ characterizes the disturbance arising from the phase and amplitude errors in the array.

The CCM in (1) can be rewritten as a sum of Kronecker products, i.e.,

$$\mathbf{R}_c = \sum_{i=1}^{N_c} \mathbf{R}_{ai} \otimes \mathbf{R}_{bi} \quad (2)$$

where $\{\mathbf{R}_{ai}\}$ are $N \times N$ the spatial covariance matrices and $\{\mathbf{R}_{bi}\}$ are $M \times M$ the temporal covariance matrices. Both $\{\mathbf{R}_{ai}\}$ and $\{\mathbf{R}_{bi}\}$ are low-rank matrices [17] and they have rank one when $\mathbf{\Gamma}_i$ and $\mathbf{\Phi}_i$ are identity matrices. We further assume that the clutters are sparse in the spatiotemporal domain [21], and thus N_c is a small number. We refer to N_c $(1 \leq N_c \leq \min(N^2, M^2))$ as the separation rank.

B. CCM Estimation for STAP

In airborne radar systems, the weight vector of the fully STAP is given by

$$\mathbf{w}_{\text{opt}} = \mathbf{R}_c^{-1} \mathbf{s}_t \quad (3)$$

where $\mathbf{s}_t$ denotes the steering vector. In practice, $\mathbf{R}_c$ is unknown and is often estimated by the SCM, i.e.,

$$\mathbf{S} = \sum_{l=1}^{L} \mathbf{z}_l \mathbf{z}_l^\dagger \quad (4)$$

where $\{\mathbf{z}_l\}$ are the training samples. The number of training samples L is usually small in nonhomogeneous environments. This is particularly true in high-dimensional systems with large antenna arrays and long pulse trains, leading to the increasing demand of training data. The ML estimate of $\mathbf{R}$ can be obtained by solving the following optimization problem:

$$\hat{\mathbf{R}} = \arg\min_{\mathbf{R} \in \mathcal{S}'} \{\log\det(\mathbf{R}) + \operatorname{tr}(\mathbf{R}^{-1}\mathbf{S})\} \quad (5)$$

where $\operatorname{tr}(\cdot)$ denotes the trace and $\mathcal{S}'$ denotes the set of Kronecker product expansions defined in (2) where $\mathbf{R}_{ai}$ and $\mathbf{R}_{bi}$ are low-rank matrices. This is a nonconvex problem.

On the other hand, *a priori* knowledge about the covariance matrix may be exploited to reduce the requirement of the training samples. Assume that an approximation $\mathbf{M}$ of the CCM $\mathbf{R}$ is available. We make the assumption that $\mathbf{R}$ is in a neighborhood of $\mathbf{M}$, i.e.,

$$\|\mathbf{R} - \mathbf{M}\|_F^2 < \epsilon \quad (6)$$

where $\|\cdot\|_F^2$ defines the Frobenius norm and ϵ reflects the degree of similarity. As such, the KA ML estimate $\hat{\mathbf{R}}$ of the CCM under the Kronecker product expansion structure constraint is formulated as

$$\hat{\mathbf{R}} = \arg\min_{\mathbf{R}} \{\log\det(\mathbf{R}) + \operatorname{tr}(\mathbf{R}^{-1}\mathbf{S})\}$$
$$\text{s.t.} \begin{cases} \mathbf{R} \in \mathcal{S}' \\ \|\mathbf{R} - \mathbf{M}\|_F^2 < \epsilon. \end{cases} \quad (7)$$

III. Knowledge-Aided Permuted Singular Value Thresholding

In this section, we propose a method to find the solution of the optimization problem in (7).

A. Permuted Singular Value Thresholding

If only the structure constraint $\mathbf{R} \in \mathcal{S}'$ is considered, a nuclear norm penalization approach [19] may be applied

$$\hat{\mathbf{R}}_{\text{PSVT}} = \arg\min_{\mathbf{R}} \left\{\|\mathcal{P}(\mathbf{R}) - \mathcal{P}(\mathbf{S})\|_F^2 + \lambda\|\mathcal{P}(\mathbf{R})\|_*\right\} \quad (8)$$

where $\|\mathbf{A}\|_* = \sum_{l=1}^{\text{rank}(\mathbf{A})} |\sigma_l(\mathbf{A})|$ denotes the nuclear norm and $\sigma_l(\mathbf{A})$ is the lth largest singular value of $\mathbf{A}$. The invertible permutation operation $\mathcal{P}(\mathbf{R})$ maps the $MN \times MN$ matrix $\mathbf{R}$ to a $M^2 \times N^2$ matrix by setting the $(n-1)M + m$th row of $\mathcal{P}(\mathbf{R})$ equal to the vectorization of $\mathbf{R}(m, n)$, where $\mathbf{R}(m, n)$ denotes the (m, n)th suboldsymbolatrix of $\mathbf{R}$, i.e., the $N \times N$ suboldsymbolatrix $\mathbf{R}(m, n) = [\mathbf{R}]_{(m-1)N+1:mN,(n-1)N+1:nN}$. Fig. 1 illustrates this permutation operator. We can now rearrange the CCM as [22]

$$\mathcal{P}(\mathbf{R}_c) = \sum_{i=1}^{N_c} \operatorname{vec}(\mathbf{R}_{ai})\operatorname{vec}(\mathbf{R}_{bi})^T. \quad (9)$$

As such, the sum of Kronecker products' structure (2) turns into a sum of vector outer products (9). In addition, a matrix of the structure in (9) can be mapped into (2) using the inverse permutation operation $\mathcal{P}^{-1}(\cdot)$.

The rearrangement operation is Frobenius norm-invariant, i.e., $\|\mathcal{P}(\mathbf{R})\|_F^2 = \|\mathbf{R}\|_F^2$. The solution of (8) has a closed form that is given by the PSVT operator as

$$\begin{aligned} \mathcal{P}\{\hat{\mathbf{R}}_{\text{PSVT}}\} &= \text{SVT}_\lambda(\mathcal{P}(\mathbf{S})) \\ &= \mathbf{U}(\operatorname{diag}(\sigma_1, \ldots, \sigma_{\min(M^2,N^2)}) - \lambda\mathbf{I})_+ \mathbf{V}^\dagger \end{aligned} \quad (10)$$

where $\mathbf{U}(\operatorname{diag}(\sigma_1, \ldots, \sigma_{\min(M^2,N^2)}))_+ \mathbf{V}^\dagger$ is the singular value decomposition of $\mathcal{P}(\mathbf{S})$ and $(\cdot)_+ \triangleq \max(\cdot, 0)$.

B. Knowledge-Aided Estimation

Now consider the KA estimation. The optimization problem is reformulated as

$$\begin{aligned} \hat{\mathbf{R}} = \arg\min_{\mathbf{R}} \Big\{ &\log\det(\mathbf{R}) + \operatorname{tr}(\mathbf{R}^{-1}\mathbf{S}) \\ &+ \lambda_1\|\mathbf{R} - \mathbf{M}\|_F^2 + \lambda_2\|\mathcal{P}(\mathbf{R})\|_*\Big\} \end{aligned} \quad (11)$$

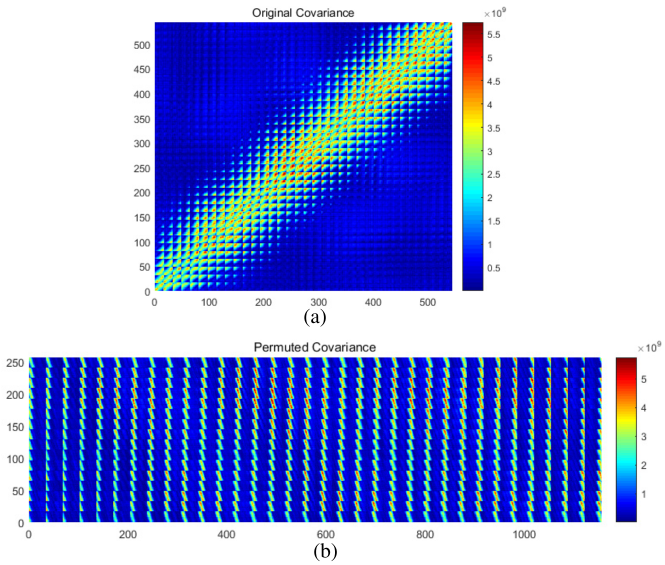

Fig. 1. Illustration of the permutation operation. (a) Original covariance matrix of size 544×544. (b) Transpose of the permuted matrix of size 1156×256.

where λ_1 is the Lagrange multiplier to account for the constraint $\|\mathbf{R}-\mathbf{M}\|_F^2 < \epsilon$. As λ_1 increases, $\mathbf{R}$ is closer to $\mathbf{M}$. In particular, as $\lambda_1 \to \infty$, the solution is $\hat{\mathbf{R}} = \mathbf{M}$. From (10), a smaller λ_2 may lead to a higher rank solution to $\mathcal{P}\{\hat{\mathbf{R}}_{\text{PSVT}}\}$, which may keep more principal components of $\mathbf{R}$ but is also more prone to the influence of noise. If $\lambda_2 \to 0$, the solution to (11) can be approximated as

$$\hat{\mathbf{R}}_{\text{KACL}} = \frac{\mathbf{S}+\lambda_1\mathbf{M}}{1+\lambda_1} \tag{12}$$

which is the same as the KACL CCM. In this letter, we set λ_2 equal to the noise power to suppress the influence of noise.

The problem of minimizing $\log\det(\mathbf{R}) + \text{tr}(\mathbf{R}^{-1}\mathbf{S})$ is not convex. An approach for tackling this issue is via relaxation [20], i.e.,

$$\hat{\mathbf{R}} = \arg\min_{\mathbf{R}} \|\mathbf{R}-\mathbf{S}\|_{\mathbf{W}}^2 \tag{13}$$

where the weighted norm with the weighting matrix $\mathbf{W}$ is defined as

$$\|\mathbf{X}\|_{\mathbf{W}} \triangleq \sqrt{\text{tr}(\mathbf{X}^\dagger\mathbf{W}^{-1}\mathbf{X})}. \tag{14}$$

We can further rewrite

$$\|\mathbf{R}-\mathbf{S}\|_{\mathbf{W}}^2 \triangleq \|\mathbf{Hr}-\mathbf{s}\|_{\mathbf{W}}^2 \triangleq (\mathbf{Hr}-\mathbf{s})^\dagger\mathbf{W}^{-1}(\mathbf{Hr}-\mathbf{s}) \tag{15}$$

where $\mathbf{s} = \text{vec}(\mathbf{S})$, and the vectorization of $\mathbf{R}$ is expressed in terms of a basis matrix $\mathbf{H}$ and the weights $\mathbf{r}$ following [23], i.e., $\text{vec}(\mathbf{R}) = \mathbf{Hr}$. The solution to (13) is then found as $\mathbf{H}\hat{\mathbf{r}}$, where

$$\hat{\mathbf{r}} = (\mathbf{W}^{-(1/2)}\mathbf{H})^{\perp}\mathbf{W}^{-(1/2)}\mathbf{s} \tag{16}$$

where $(\cdot)^{\perp}$ denotes the matrix pseudoinverse. If the weighting matrix is chosen as $\mathbf{W} = \mathbf{R}^T \otimes \mathbf{R}$, the resulting estimator can achieve the Cramér–Rao bound [12]. Since $\mathbf{R}$ is unknown, we use the SCM to construct the weighting matrix as $\mathbf{W}_S = \mathbf{S}^T \otimes \mathbf{S}$.

Now the relaxed problem is formulated as

$$\hat{\mathbf{R}} = \arg\min_{\mathbf{R}} \Big\{ \|\mathbf{R}-\mathbf{S}\|_{\mathbf{W}_S}^2 + \lambda_1\|\mathbf{R}-\mathbf{M}\|_F^2 + \lambda_2\|\mathcal{P}(\mathbf{R})\|_* \Big\}. \tag{17}$$

Exploiting the proximal algorithm [24], we propose Algorithm 1 to find the global minimizer of (17). In Algorithm 1, τ_k denotes the step size and can be fixed to a constant $\tau < 1$. ε is a value small enough to guarantee convergence. $\{\mathbf{U}_i\}$ denotes the running sum of the errors. When the algorithm converges, $\{\mathbf{U}_i\}$ converges, and $\bar{\mathbf{R}} = \mathbf{R}_1 = \mathbf{R}_2 = \mathbf{R}_3$. In Step 3, the minimization problem is solved following the solution to (13) with $\mathbf{H} = \mathbf{I}$, assuming that the CCM is Hermitian. The PSVT operation in Step 5 guarantees the Kronecker product structure.

Algorithm 1 Proximal Gradient KAPSVT

1: Initialize $\mathbf{U}_1^0$, $\mathbf{U}_2^0$, $\mathbf{U}_3^0$, and $\bar{\mathbf{R}}^0$ to be all-zero matrices. Choose step sizes τ_k, where k is the number of iterations. Initialize $k = 0$.
2: **while** $\|\bar{\mathbf{R}}^{k+1} - \bar{\mathbf{R}}^k\|_F^2 < \varepsilon$ **do**
3:

$$\mathbf{R}_1^{k+1} = \arg\min_{\mathbf{R}} \left\|(1+\tau_k)\mathbf{R} - \mathbf{S} - \tau_k(\bar{\mathbf{R}}^k - \mathbf{U}_1^k)\right\|_{\mathbf{W}_S}^2$$

4:

$$\mathbf{R}_2^{k+1} = \frac{\lambda_1\mathbf{M} + \tau_k(\bar{\mathbf{R}}^k - \mathbf{U}_2^k)}{\lambda_1 + \tau_k}$$

5:

$$\mathbf{R}_3^{k+1} = \mathcal{P}^{-1}\left\{\text{SVT}_{\frac{\lambda_2}{\tau_k}}\left[\mathcal{P}(\bar{\mathbf{R}}^k) - \mathcal{P}(\bar{\mathbf{U}}_3^k)\right]\right\}$$

6: Set

$$\bar{\mathbf{R}}^{k+1} = \frac{1}{3}\left(\mathbf{R}_1^{k+1} + \mathbf{R}_2^{k+1} + \mathbf{R}_3^{k+1}\right)$$
$$\mathbf{U}_i^{k+1} = \mathbf{U}_i^k + \mathbf{R}_i^{k+1} - \bar{\mathbf{R}}^{k+1}, \quad i = 1,2,3$$

7: **end while**
8: **return** $\hat{\mathbf{R}}_{\text{KAPSVT}} = \bar{\mathbf{R}}^{k+1}$

IV. Experiments and Analysis

In this section, the proposed KAPSVT estimator is compared with the KACL, PSVT, and SCM estimators for real data generated from an airborne radar experiment, with the parameters listed in Table I. The experiment was conducted around the mountains on the edge of the Yellow River in Shaanxi Province on October 13, 2014. The radar system is equipped with half-wavelength spaced elements and operated in the side-looking mode at a speed of 120 m/s. The priori matrix $\mathbf{M}$ used for the KA CCM estimation is set as the CCM estimated using SCM from 1000 samples obtained from a previous experiment using the same airborne radar, flight path, and velocity on the same day.

Fig. 2 demonstrates the spatial–temporal clutter spectrum of the CUT, generated from the SCM estimate of the CCM using 1000 homogenous training samples, which can be regarded as the theoretical CCM according to [17]. Fig. 3 presents the

TABLE I
PARAMETER LIST

Parameter	Value	Parameter	Value
Pulse repeat interval	0.004s	Carrier frequency	448MHz
Azimuth angle	0	M	34
Elevation angel	0.0109	N	16

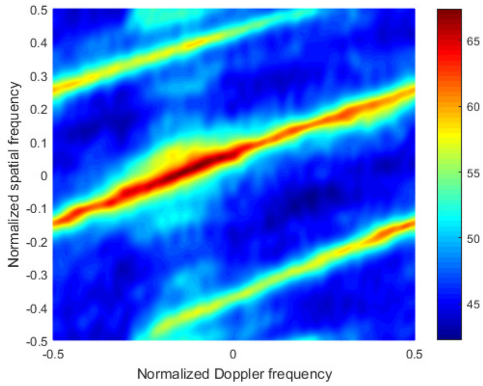

Fig. 2. Spatiotemporal clutter spectrum of the CUT with a slope of $\beta = 2.82$.

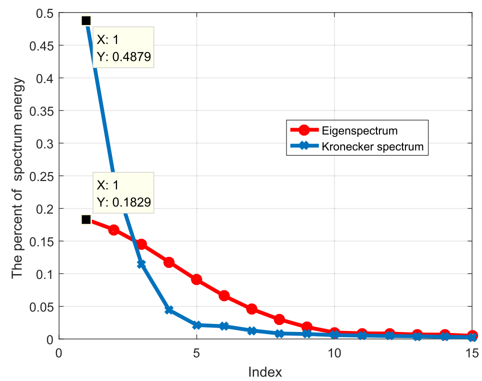

Fig. 3. Comparison of the percent of spectrum energy of eigenspectrum and Kronecker spectrum. The first component of Kronecker spectrum contains 48.79% of the spectrum energy, and the first component of eigenspetrum contains 18.29% energy.

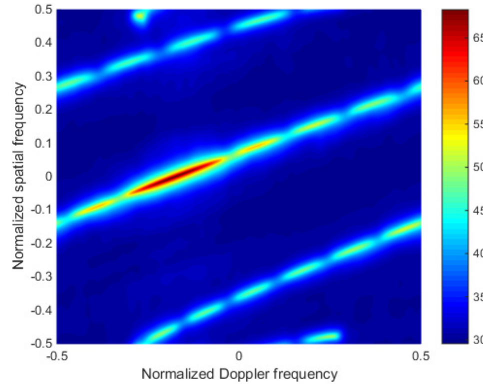

Fig. 4. Spatiotemporal clutter spectrum of estimated CCM by KAPSVT with $L = 200$, $\tau = 1$, $\lambda_1 = 0.1$, and λ_2 determined by the noise power.

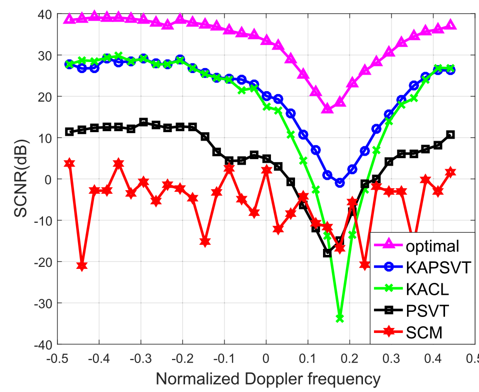

Fig. 5. SCNR performance of the proposed algorithm compared with other methods with the number of training data $L = 200$.

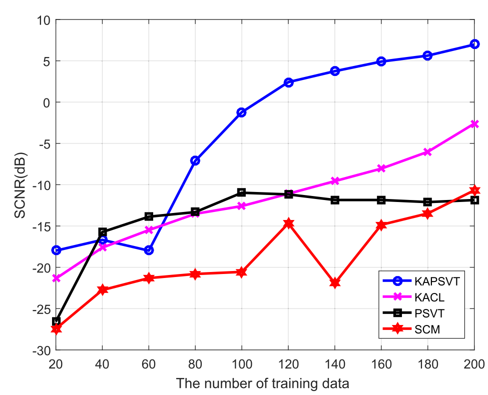

Fig. 6. SCNR performance as a function of the number of training data.

eigenspectra for the CUT, i.e., the eigenvalues of $\mathbf{R}_c$ of the CUT, and the Kronecker spectrum, i.e., the eigenspectrum of $\mathcal{P}(\mathbf{R}_c)$. It is seen that the energy is more concentrated in the principal eigenvalues for the Kronecker spectrum compared with the eigenspectrum. This suggests that the permuted CCM can be represented using fewer principal components and thus may be estimated with less training data.

Fig. 4 demonstrates the spatiotemporal spectrum of the CCM estimated by the proposed KAPSVT estimator, with $L = 200$, $\tau = 1$, and $\lambda_1 = 0.1$. It is seen that the estimated spectrum in Fig. 4 agrees well with the theoretical results in Fig. 2.

We also use the signal-to-clutter-noise ratio (SCNR) of the filter output to assess the quality of the estimated CCM, which is defined as

$$\mathrm{SCNR} = \frac{|\hat{\mathbf{w}}^{\dagger}\mathbf{s}_t|^2}{\hat{\mathbf{w}}^{\dagger}\mathbf{R}_c\hat{\mathbf{w}}} \tag{18}$$

where $\hat{\mathbf{w}}$ is the estimated filter constructed using the estimated CCM. The clutter-to-noise ratio is about 20 dB in the experiments. Fig. 5 shows the SCNR for the optimal estimator, the proposed estimator, the KACL, PSVT, and SCM estimators. The results show that the KA algorithms significantly outperform alternative estimators. The proposed KAPSVT estimator, which exploits the structure of CCM, can further improve the SCNR, especially for low-speed regimes.

Fig. 6 shows the SCNR performance versus the number of training data. The proposed estimator generally outperforms

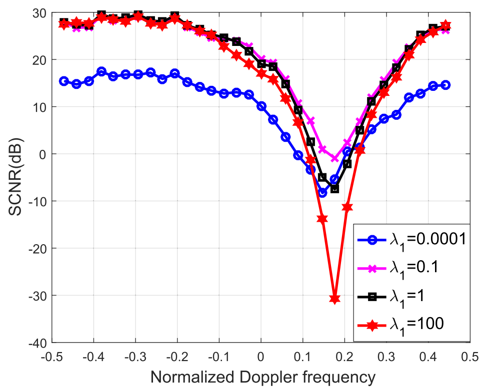

Fig. 7. SCNR performance of the proposed algorithms under different λ_1 with $\tau_k = 0.1$, $L = 200$, and λ_2 determined by the noise power.

other estimators at different L. The parameter λ_1 plays an important role in the proposed algorithm. The performance with different λ_1 is depicted in Fig. 7. The SCNR with the KAPSVT estimator is approximate to the SCNR with the PSVT estimator when λ_1 approaches zero. If λ_1 is larger, e.g., $\lambda_1 = 100$, the *a priori* matrix $\mathbf{M}$ plays a more significant role for the KA estimator and the current training data are less exploited, which can also result in significant performance loss. Automatically choosing λ_1 is left for future work.

V. Conclusion

In this letter, we study the estimation of the structured CCM in airborne STAP, which is a sum of low-rank Kronecker products. The proposed KAPSVT method incorporates the training data, KA covariance matrix, and structure of the CCM into the ML estimation, and a proximal algorithm is designed to find the estimator. An invertible permutation operation is exploited to transform the Kronecker product constraint into a linear constraint. It has been shown by real data experiments that the CCM estimated by the KAPSVT effectively improves the SCNR performance for airborne radar systems.

References

[1] Z. Yang, X. Li, H. Wang, and W. Jiang, "On clutter sparsity analysis in space–time adaptive processing airborne radar," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 5, pp. 1214–1218, Sep. 2013.

[2] Z. Wang, Y. Wang, K. Duan, and W. Xie, "Subspace-augmented clutter suppression technique for STAP radar," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 3, pp. 462–466, Mar. 2016.

[3] Y. Wu, T. Wang, J. Wu, and J. Duan, "Training sample selection for space-time adaptive processing in heterogeneous environments," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 4, pp. 691–695, Apr. 2015.

[4] A. Liu, H. Sun, K. C. Teh, C. J. Baker, and C. Gao, "Robust space-time adaptive processing for nonhomogeneous clutter in the presence of model errors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 1, pp. 155–168, Feb. 2016.

[5] X. Zhang, Z. He, B. Liao, X. Zhang, Z. Cheng, and Y. Lu, "A^{2}RC: An accurate array response control algorithm for pattern synthesis," *IEEE Trans. Signal Process.*, vol. 65, no. 7, pp. 1810–1824, Apr. 2017.

[6] X. Zhang, Z. He, B. Liao, X. Zhang, and W. Peng, "Pattern synthesis with multipoint accurate array response control," *IEEE Trans. Antennas Propag.*, vol. 65, no. 8, pp. 4075–4088, Aug. 2017.

[7] J. R. Guerci and E. J. Baranoski, "Knowledge-aided adaptive radar at DARPA: An overview," *IEEE Signal Process. Mag.*, vol. 23, no. 1, pp. 41–50, Jan. 2006.

[8] O. Besson, J. Y. Tourneret, and S. Bidon, "Knowledge-aided Bayesian detection in heterogeneous environments," *IEEE Signal Process. Lett.*, vol. 14, no. 5, pp. 355–358, May 2007.

[9] Y. Gao, H. Li, and B. Himed, "Knowledge-aided range-spread target detection for distributed MIMO radar in nonhomogeneous environments," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 617–627, Feb. 2017.

[10] S. Zhu, G. Liao, J. Xu, and Y. Gao, "Robust space–time adaptive processing with colored loading using iterative optimization," *Digit. Signal Process.*, vol. 35, pp. 14–20, Dec. 2014.

[11] S. Zhang, Z. He, J. Li, and Y. Wang, "A robust colored-loading factor optimization approach for knowledge-aided STAP," in *Proc. IEEE Radar Conf.*, Philadelphia, PA, USA, May 2016, pp. 1–5.

[12] Y. Wang and Z. He, "Thinned knowledge-aided STAP by exploiting structural covariance matrix," *IET Radar, Sonar Navigat.*, vol. 11, no. 8, pp. 1266–1275, Jun. 2017.

[13] R. S. Raghavan, "CFAR detection in clutter with a Kronecker covariance structure," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 2, pp. 619–629, Apr. 2017.

[14] B. Kang. "Robust covariance matrix estimation for radar space-time adaptive processing (STAP)," unpublished paper, 2016. [Online]. Available: https://arxiv.org/abs/1602.05069

[15] T. Tsiligkaridis, A. O. Hero, III, and S. Zhou, "On convergence of Kronecker graphical lasso algorithms," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1743–1755, Apr. 2013.

[16] G. Sun, Z. He, F. Jia, and R. Li, "Kronecker product PCA for structured covariance matrix of airborne radar STAP," in *Proc. IEEE Radar Conf.*, Seattle, WA, USA, May 2017, pp. 1015–1019.

[17] J. Ward, "Space-time adaptive processing for airborne radar," MIT Lincoln Lab., Lexington, MA, USA, Tech. Rep. 1015, Dec. 1994.

[18] K. Greenewald and A. O. Hero, III, "Robust Kronecker product PCA for spatio-temporal covariance estimation," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6368–6378, Dec. 2015.

[19] T. Tsiligkaridis and A. O. Hero, III, "Covariance estimation in high dimensions via Kronecker product expansions," *IEEE Trans. Signal Process.*, vol. 61, no. 21, pp. 5347–5360, Nov. 2013.

[20] B. Ottersten, P. Stoica, and R. Roy, "Covariance matching estimation techniques for array signal processing applications," *Digit. Signal Process.*, vol. 8, no. 3, pp. 185–210, Jul. 1998.

[21] Z. Yang, X. Li, H. Wang, and R. Fa, "Knowledge-aided STAP with sparse-recovery by exploiting spatio-temporal sparsity," *IET Signal Process.*, vol. 10, no. 2, pp. 150–161, 2016.

[22] C. F. Van Loan and N. Pitsianis, "Approximation with Kronecker products," in *Linear Algebra for Large Scale and Real-Time Applications*. Dordrecht, The Netherlands: Springer, 1993, pp. 293–314.

[23] I. Soloveychik and A. Wiesel, "Tyler's covariance matrix estimator in elliptical models with convex structure," *IEEE Trans. Signal Process.*, vol. 62, no. 20, pp. 5251–5259, Oct. 2014.

[24] N. Parikh and S. Boyd, "Proximal algorithms," *Found. Trends Optim.*, vol. 1, no. 3, pp. 127–239, Jan. 2014.