ALGORITHMS AND PERFORMANCE ANALYSIS FOR ESTIMATION OF LOW-RANK MATRICES WITH KRONECKER STRUCTURED SINGULAR VECTORS

Raj Tejas Suryaprakash, Brian E. Moore and Raj Rao Nadakuditi

Dept. of EECS, University of Michigan, Ann Arbor, USA

ABSTRACT

We consider the problem of estimating the singular vectors of low-rank signal matrices buried in noise in the setting where the singular vectors are assumed to be Kronecker products of unknown vectors. We propose four algorithms for estimating such singular vectors, analyze their performance and show that they asymptotically fail to estimate to latent singular vector below the same critical SNR. We corroborate our theoretical findings with numerical simulations and illustrate the improved performance on a STAP beamforming application.

Index Terms— Kronecker products, Random Matrices, Singular Value Decomposition, Space Time Adaptive Processing

1. INTRODUCTION

The singular value decomposition (SVD) is used to estimate low-rank signal matrices buried in noise. Here we consider the setting where the left singular vectors of the low-rank signal matrices are Kronecker structured, *i.e.*, they are Kronecker products of some unknown unit-norm vectors. In this setting, we might expect to improve the estimation of the low-rank signal matrices by exploiting the underlying Kronecker structure.

In recent work, Tsiligkaridis and Hero [1] considered the problem of estimation of (full-rank) covariance matrices which were Kronecker products of (full-rank) covariance matrices. They developed an algorithm that builds on the seminal work of Pitsianis and Van Loan [2] on the estimation of matrices that can be expressed as sums of matrices having Kronecker product structure. A rearrangement operator plays an important role in their algorithm since it rearranges the data to exploit the redundancy in the data due to the Kronecker product. In their algorithm, the rearrangement of the data is done first and then an SVD is applied on the rearranged data.

In this paper, we consider a different model than that considered by Tsiligkaridis and Hero in [1]. Here we consider a signal-plus-model where the latent low-rank signal-plus-noise matrix is modeled as a Kronecker product of unstructured low-rank matrices. We propose four algorithms for estimating the Kronecker structured singular vector that utilizes the rearrangement operator in different ways. We characterize the estimation performance of the associated singular vector estimate and bring into sharp focus a phase transition phenomenon which separates a regime where the estimated singular vector is correlated with the latent singular vector from a regime where the estimated singular vector is orthogonal to the latent singular vector. Surprisingly, all four of the proposed algorithms fail at *exactly the same SNR*.

Our performance analysis also reveals the best performance is achieved by the algorithm which first computes the SVD and then utilizes the Kronecker structure to effectively denoise the estimate. Rearranging the data first to exploit the Kronecker structure results in a performance loss.

The development of the algorithms and their associated performance analysis is the primary contribution of this paper, which is organized as follows. We describe the signal-plus-noise model in Section 2, the rearrangement operator in Section 3 and summarize the algorithms in Section 4. We analyze the performance in Section 5, illustrate it with numerical simulations in Section 6 and offer some concluding remarks in Section 7.

2. SYSTEM MODEL

Consider the signal-plus-noise model

$$X = \sigma_0 \underbrace{(a \otimes b)}_{:=u} s^H + Z, \tag{1}$$

where $m = pq$, $X \in \mathbb{C}^{m \times n}$ is the observed data matrix, $\sigma_0 \in \mathbb{R}_+$ is the signal-to-noise ratio (SNR). The vectors $a \in \mathbb{C}^{p \times 1}$ and $b \in \mathbb{C}^{q \times 1}$ are unknown, non-random unit norm vectors and $u = a \otimes b \in \mathbb{C}^{m \times 1}$ is the unit-norm signal subspace basis vector, where $\otimes$ is the Kronecker product [3, Chapter 13]. The vector $s \in \mathbb{C}^{n \times 1} \sim \mathcal{CN}(0, I_n)$ and Z is the noise-only matrix with i.i.d. zero mean, unit variance entries with bounded higher order moments. Higher rank extensions of the model in (1) are being investigated; their analysis is beyond the scope of this paper. The model in (1) is motivated by applications such as Moving Target Indicator (MTI) radar system, where the space-time manifold vector associated with clutter at normalized angular location θ_c has precisely the form [4, 5]

$$u(\theta_c, f_c) = a(f_c) \otimes b(\theta_c), \tag{2}$$

where

$$b(\theta_c) = \frac{1}{\sqrt{p}} \begin{bmatrix} 1 & e^{j\,2\pi\,\theta_c} & \dots & e^{j\,2\pi\,(p-1)\,\theta_c} \end{bmatrix}^T \tag{3}$$

and

$$a(f_c) = \frac{1}{\sqrt{q}} \begin{bmatrix} 1 & e^{j\,2\pi\,f_d} & \dots & e^{j\,2\pi\,(q-1)\,f_d} \end{bmatrix}^T. \tag{4}$$

In (4), f_d is the normalized Doppler shift induced by clutter relative to the array boresight.

In the MTI application, the Space-Time Adaptive Processing (STAP) filter [6], assuming σ_0 is known, is computed as [7]

$$\widehat{w} = (I + \sigma_0^2 \widehat{u}\,\widehat{u}^H)^{-1}\, u(\theta_t, f_t), \tag{5}$$

Work supported by ONR Young Investigator Award N00014-11-1-0660, US Army Research Office (ARO) under grant W911NF-11-1-0391, NSF grant CCF-1116115, and AFOSR Young Investigator Award FA9550-12-1-0266.

where $u(\theta_t, f_t)$ is the space-time vector associated with the target, and has the same form as (2) and $\widehat{u}$ is an estimate of u.

The left singular vector associated with the largest singular value in the SVD of X yields an estimate $\widehat{u}_{\rm svd}$ of u which does not exploit the Kronecker structure of the latent left singular vector. Our objective is to develop algorithm(s) that exploit the Kronecker structure of the left singular vectors in (1) to improve the estimation of the vector u relative to $\widehat{u}_{\rm svd}$ and to characterize the associated performance loss(es) due to limited, noisy data.

3. REARRANGEMENT OPERATOR FOR KRONECKER DECOMPOSITION

In this section, we define the rearrangement operator which will be used extensively in the algorithms and is key to exploiting the Kronecker structure. Assume that we have the matrix $A \in \mathbb{C}^{pq\times m_1 m_2} = B \otimes C$, where $B \in \mathbb{C}^{p\times m_1}$ and $C \in \mathbb{C}^{q\times m_2}$. Furthermore, assume that A has a uniform blocking given by

$$A = \begin{bmatrix} A_{11} & \dots & A_{1,m_1} \\ \vdots & \ddots & \vdots \\ A_{p,1} & \dots & A_{p,m_1} \end{bmatrix}, \quad A_{ij} \in \mathbb{C}^{q\times m_2} \tag{6}$$

we have that [2]

$$||A - B\otimes C||_F = ||\mathcal{R}(A) - \mathrm{vec}(B)\,(\mathrm{vec}(C))^T\,||_F, \tag{7}$$

where the $\mathrm{vec}(\cdot)$ operation stacks the columns of its matrix argument into a vector:

$$M \in \mathbb{C}^{l\times m} \implies \mathrm{vec}(M) = \begin{bmatrix} M(1:l,1) & \dots & M(1:l,m) \end{bmatrix}^T, \tag{8}$$

and

$$\mathcal{R}(A) = \begin{bmatrix} A_1 \\ \vdots \\ A_{m_1} \end{bmatrix}, \quad A_j = \begin{bmatrix} \mathrm{vec}(A_{1,j})^T \\ \vdots \\ \mathrm{vec}(A_{q,j})^T \end{bmatrix}, \quad j = 1:m_1. \tag{9}$$

We will utilize (7) extensively in what follows.

4. NEW ALGORITHMS

Figure 1 summarizes the algorithms used to obtain $\widehat{u}$. We denote the estimates $\widehat{u}_{\rm svd}$, $\widehat{u}_{\rm kpca,1}$, $\widehat{u}_{\rm kpca,2\ell}$, $\widehat{u}_{\rm kpca,2r}$ and $\widehat{u}_{\rm kpca,3}$, since they are based on the SVD, and four versions of 'Kronecker PCA's of X, respectively. Note that the rearrangement operator features prominently in the algorithms. $\widehat{u}_{\rm svd}$ is the left singular value of X corresponding to its largest singular value. Due to space constraints, we describe only one Kronecker PCA algorithm ($\mathrm{kpca}, 2\ell$). The remaining algorithms may be interpreted using Figure 1 similarly.

In step 1 of Algorithm $\mathrm{kpca}, 2\ell$, we apply the rearrangement operator defined in (9) to X given by (1). Let the dimensions of the rearrangement be $\mathcal{R}(X) : \mathbb{C}^{pq\times n} \to \mathbb{C}^{p\times qn}$. Upon rearrangement,

$$X_{\mathcal{R}} := \mathcal{R}(X) = \sigma\, a\, \mathrm{vec}\left(bs^H\right)^T + \mathcal{R}(Z), \tag{10}$$

where $\mathrm{vec}(\cdot)$ operation is defined in (8), and we have used the fact that $\mathrm{vec}(a) = a$. Let

$$X_{\mathcal{R}} = \sigma_{\mathcal{R}}\, \widehat{a}\, \widehat{v}^H \tag{11}$$

be the SVD of $X_{\mathcal{R}}$ truncated to rank 1. Comparing (10) and (11), we observe that we can use $\widehat{a}$ to estimate a and the complex conjugate of $\widehat{v}$, denoted $\widehat{v}^*$ to estimate $\mathrm{vec}(bs^H)$. In step 2, we return $\widehat{a}_{\rm kpca,2\ell} = \widehat{a}$ in (11) as the estimate of a. In steps 3 and 4, we reshape $\widehat{v}^*$: since $\widehat{v}^* \in \mathcal{C}^{qn}$ is an estimate of $\mathrm{vec}(bs^H)$, we may reshape it as

$$\mathrm{reshape}_{qn\to q\times n}(\widehat{v}^*) = \begin{bmatrix} \widehat{v}^*(1:q) & \dots & \widehat{v}^*((n-1)q+1:nq) \end{bmatrix}, \tag{12}$$

so that $\widehat{V} := \mathrm{reshape}_{qn\to q\times n}(\widehat{v}^*)$ is an estimate of the matrix bs^H. In other words,

$$\widehat{V} = bs^H + \Delta_{\widehat{V}}, \tag{13}$$

where $\Delta_{\widehat{V}}$ is the error in estimating bs^H. As suggested by (13), in step 5, we return $\widehat{b}_{\rm kpca,2\ell}$ = the left singular vector of $\widehat{V}$ corresponding to its largest singular value, as the estimate of b. Finally, in step 6, we exploit the Kronecker structure of u in (1) and return as its estimate

$$\widehat{u}_{\rm kpca,2\ell} = \widehat{a}_{\rm kpca,2\ell} \otimes \widehat{b}_{\rm kpca,2\ell},$$

where $\widehat{a}_{\rm kpca,2\ell}$ was estimated in step 2 and $\widehat{b}_{\rm kpca,2\ell}$ in step 5 as described above. The other algorithms can be derived in a similar manner. Note that in $\mathrm{kpca}, 1$ we first compute $\widehat{u}_{\rm svd}$ and then apply the SVD to $\mathcal{R}(\widehat{u}_{\rm svd})$. The other algorithms apply the rearrangement operator first and then compute the SVD(s).

5. PERFORMANCE ANALYSIS & LIMITS

We now provide a unified performance characterization of the various KPCA algorithms.

Claim 1. *Let X be modeled as in* (1)*. Let $\widehat{u}_{\rm svd}$, $\widehat{u}_{\rm kpca,1}$, $\widehat{u}_{\rm kpca,2\ell}$, $\widehat{u}_{\rm kpca,2r}$, $\widehat{u}_{\rm kpca,3}$ be estimates of $u = a\otimes b$ computed via the algorithms in Figure 1. Then, we have that,*

$$|\langle \widehat{u}_{\rm svd}, u\rangle|^2 \xrightarrow{a.s.} \alpha^2(\sigma_0, c_{\rm svd}), \tag{14a}$$

$$|\langle \widehat{u}_{\rm kpca,1}, u\rangle|^2 \xrightarrow{a.s.} \alpha^2(\sigma_{\rm a,1}, c_{\rm a,1})\,\beta^2(\sigma_{\rm b,1}, c_{\rm b,1}) \tag{14b}$$

$$|\langle \widehat{u}_{\rm kpca,2\ell}, u\rangle|^2 \xrightarrow{a.s.} \alpha^2(\sigma_{\rm a,2\ell}, c_{\rm a,2\ell})\,\alpha^2(\sigma_{\rm b,2\ell}, c_{\rm b,2\ell}) \tag{14c}$$

$$|\langle \widehat{u}_{\rm kpca,2r}, u\rangle|^2 \xrightarrow{a.s.} \alpha^2(\sigma_{\rm a,2r}, c_{\rm a,2r})\,\beta^2(\sigma_{\rm b,2r}, c_{\rm b,2r}) \tag{14d}$$

$$|\langle \widehat{u}_{\rm kpca,3}, u\rangle|^2 \xrightarrow{a.s.} \alpha^2(\sigma_{\rm a,3}, c_{\rm a,3})\,\beta^2(\sigma_{\rm b,3}, c_{\rm b,3}), \tag{14e}$$

as $p, q, n \to \infty$, $pq/n \to c_{\rm svd}$, $p/q \to c_{\rm a,1}(= c_{\rm b,1})$, $p/(qn) \to c_{\rm a,2\ell}(= c_{\rm a,3})$, $pn/q \to c_{\rm b,2r}(= c_{\rm b,3})$, $q/n \to c_{\rm b,2\ell}$ and $p/n \to c_{\rm a,2r}$. Here,

$$\alpha^2(\sigma, c) := \begin{cases} 1 - \dfrac{c(1+\sigma^2)}{\sigma^2(\sigma^2+c)} & \text{if } \sigma_0 > c_{\rm svd}^{1/4} \\ 0 & \text{otherwise} \end{cases}, \tag{15a}$$

$$\beta^2(\sigma, c) := \begin{cases} 1 - \dfrac{(c+\sigma^2)}{\sigma^2(\sigma^2+1)} & \text{if } \sigma_0 > c_{\rm svd}^{1/4} \\ 0 & \text{otherwise} \end{cases}. \tag{15b}$$

The quantities $\sigma_{\rm svd}, \dots, c_{\rm b,3}$ in the arguments of the right hand side (14) *are given by the Table 1, in which*

$$\alpha^2_{\rm svd} := \alpha^2(\sigma_0, c_{\rm svd}), \qquad \beta^2_{\rm svd} := \beta^2(\sigma_0, c_{\rm svd}).$$

Justification. Eq. (14a) may be shown rigorously by direct application of [8, Theorem 2.10]. Eq. (14e) may be shown rigorously by application of [8, Theorem 2.10] twice: to steps 2 and 4 of algorithm 5, which yield characterizations of $|\langle \widehat{a}_{\rm kpca,3}, a\rangle|^2$ and $|\langle \widehat{b}_{\rm kpca,3}, b\rangle|^2$, and noting that $(\widehat{a}_{\rm kpca,3}\otimes\widehat{b}_{\rm kpca,3})^H(a\otimes b) = (\widehat{a}^H_{\rm kpca,3}a)\,(\widehat{b}^H_{\rm kpca,3}b)$. Equations (14b) to (14d) involve proof via assuming that the sample singular vectors have a signal+noise structure, and will be presented in the journal version of this paper. □

Alg. No.	Description	Algorithm	Performance Terms
1	Singular Value Decomposition	1. $\widehat{u}_{\rm svd}$ = largest left singular vector of X	σ_0, $c_{\rm svd} = pq/n$
2	Kronecker PCA (Version 1) $({\rm kpca}, 1)$	1. $\widehat{u}_{\rm svd}$ = largest left singular vector of X 2. $U_{{\rm svd},\mathcal{R}} = \mathcal{R}_{pq\times 1\rightarrow p\times q}(\widehat{u}_{\rm svd})$ 3. $\widehat{a}_{{\rm kpca},1}$ = largest left singular vector of $U_{{\rm svd},\mathcal{R}}$ 4. $\widehat{b}_{{\rm kpca},1}$ = (largest right singular vector of $U_{{\rm svd},\mathcal{R}})^*$ 5. $\widehat{u}_{{\rm kpca},1} = \widehat{a}_{{\rm kpca},1} \otimes \widehat{b}_{{\rm kpca},1}$	$\sigma_{{\rm a},1} = \sqrt{\alpha^2_{\rm svd}\, p}/\sqrt{1-\alpha^2_{\rm svd}}$ $c_{{\rm a},1} = p/q$ $\sigma_{{\rm b},1} = \sqrt{\alpha^2_{\rm svd}\, p}/\sqrt{1-\alpha^2_{\rm svd}}$ $c_{{\rm b},1} = p/q$
3	Kronecker PCA (Version 2ℓ) $({\rm kpca}, 2\ell)$	1. $X_{\mathcal{R}} = \mathcal{R}_{pq\times n\rightarrow p\times qn}(X)$ 2. $\widehat{a}_{{\rm kpca},2\ell}$ = largest left singular vector of $X_{\mathcal{R}}$ 3. $\widehat{v}$ = largest right singular vector of $X_{\mathcal{R}}$ 4. $\widehat{V} = {\rm reshape}_{qn\rightarrow q\times n}(\widehat{v}^*)$ 5. $\widehat{b}_{{\rm kpca},2\ell}$ = largest left singular vector of $\widehat{V}$ 6. $\widehat{u}_{{\rm kpca},2\ell} = \widehat{a}_{{\rm kpca},2\ell} \otimes \widehat{b}_{{\rm kpca},2\ell}$	$\sigma_{{\rm a},2\ell} = \sigma_0/\sqrt{q}$ $c_{{\rm a},2\ell} = p/(qn)$ $\sigma_{{\rm b},2\ell} = \sqrt{\beta^2_{\rm svd}\, q}/\sqrt{1-\beta^2_{\rm svd}}$ $c_{{\rm b},2\ell} = q/n$
4	Kronecker PCA (Version $2r$) $({\rm kpca}, 2r)$	1. $X_{\mathcal{R}'} = \mathcal{R}_{pq\times n\rightarrow pn\times q}(X)$ 2. $\widehat{b}_{{\rm kpca},2r}$ = (largest right singular vector of $X_{\mathcal{R}'})^*$ 3. $\widehat{u}$ = largest left singular vector of $X_{\mathcal{R}'}$ 4. $\widehat{U} = {\rm reshape}_{pn\rightarrow p\times n}(\widehat{u})$ 5. $\widehat{a}_{{\rm kpca},2r}$ = largest left singular vector of $\widehat{U}$ 6. $\widehat{u}_{{\rm kpca},2r} = \widehat{a}_{{\rm kpca},2r} \otimes \widehat{b}_{{\rm kpca},2r}$	$\sigma_{{\rm a},2r} = \sqrt{\alpha^2_{\rm svd}\, p}/\sqrt{1-\alpha^2_{\rm svd}}$ $c_{{\rm a},2r} = p/n$ $\sigma_{{\rm b},2r} = \sigma_0\sqrt{n}/\sqrt{q}$ $c_{{\rm b},2r} = pn/q$
5	Kronecker PCA (Version 3) $({\rm kpca}, 3)$	1. $X_{\mathcal{R}} = \mathcal{R}_{pq\times n\rightarrow p\times qn}(X)$ 2. $\widehat{a}_{{\rm kpca},3}$ = largest left singular vector of $X_{\mathcal{R}}$ 3. $X_{\mathcal{R}'} = \mathcal{R}_{pq\times n\rightarrow pn\times q}(X)$ 4. $\widehat{b}_{{\rm kpca},3}$ = (largest right singular vector of $X_{\mathcal{R}'})^*$ 5. $\widehat{u}_{{\rm kpca},3} = \widehat{a}_{{\rm kpca},3} \otimes \widehat{b}_{{\rm kpca},3}$	$\sigma_{{\rm a},3} = \sigma_0/\sqrt{p}$ $c_{{\rm a},3} = p/(qn)$ $\sigma_{{\rm b},3} = \sigma_0\sqrt{n}/\sqrt{q}$ $c_{{\rm b},3} = pn/q$

Table 1: Algorithms for estimating u in (1). Largest left/right singular vector means the left/right singular vector corresponding to the largest singular value. $\mathcal{R}(X)$ rearranges X using the rearrangement operator defined in (9). Where required, we have made rearrangement dimensions explicit e.g., $\mathcal{R}_{pq\times n\rightarrow p\times qn}$ takes a matrix of dimensions $pq \times n$ and returns a matrix of dimensions $p \times qn$. ${\rm reshape}_{qn\rightarrow q\times n}(\cdot)$ is defined in (12). Superscript * stands for 'complex conjugate of'.

6. NUMERICAL SIMULATIONS

Figure 1 plots the inner-product-squared accuracy, $|\langle \widehat{u}, u\rangle|^2$, where $\widehat{u}$ is estimated using all the algorithms in Table 1 (i.e. $\widehat{u} = \{\widehat{u}_{\rm svd}, \widehat{u}_{{\rm kpca},1}, \widehat{u}_{{\rm kpca},2\ell}, \widehat{u}_{{\rm kpca},2r}, \widehat{u}_{{\rm kpca},3}\}$) versus SNR σ_0, for a system with $p = 15$, $q = 15$ and $n = 200$ (i.e. sample deficient regime). The empirical points were computed over 10^4 Monte Carlo simulations, by generating data via (1) and implementing algorithms as in Table 1. The theoretical curves were plotted using the theoretical expressions given by (14), and the agreement with experimental data in the plot validates these theoretical expressions. In Figure 2, we plot the logarithm of the empirical inner-product-squared accuracy for Algorithm $({\rm kpca}, 2\ell)$, over 5000 trials, for different values of SNR σ_0 and aspect ratios $c_{\rm svd} = (pq/n)$. The plot highlights a phase transition in algorithm performance which separates a regime where $\widehat{u}$ returned by the algorithms is accurate (red) from a regime where the the performance breaks down (blue). The line corresponding to $\sigma_{\rm crit}$ plots this phase transition boundary predicted by the theoretical expressions in (14), and accurately predicts the empirically observed boundary.

Finally we considered the MTI application to illustrate the improvement in performance relative to the SVD for the estimation of STAP filters, given by (5). Using (3) and (4), we simulated a system with $q = 15$ sensors in a uniform linear array spaced half wavelength ($0.5 \times 0.67m$) apart, which emits $p = 15$ pulses in one coherent pulse interval. The target was chosen to be broadside to the array ($\theta_t = 0$) and induce normalized Doppler shift $f_t = 0$. A single clutter was located at 30^o, and induced normalized Doppler shift $f_c = 0.0995$ units at the array. In this application, σ_0 in (1) has the interpretation of clutter-to-noise ratio (CNR). The goal of STAP is to suppress clutter. Figure 3 shows the improvement in clutter suppression when we use the KPCA based algorithms to estimate $\widehat{u}$ against using only the SVD, We plot the filter output power at clutter location and Doppler frequency, (θ_c, f_c) over different CNRs, and we see that all KPCA based algorithms suppress clutter more than regular SVD, by about 10 dB. In Figure 4, we plot the beampatterns of the array, for $\widehat{w}$ in (5) was derived using SVD and ${\rm kpca}, 2\ell$ and compare their performance to the optimum filter which uses the true Σ. As expected the new algorithm(s) outperform the SVD. In all simulations, we see that the $({\rm kpca}, 1)$ outperforms the other KPCA variants as predicted by the theory.

7. CONCLUSIONS

We proposed four algorithms for estimating singular vectors with Kronecker product structure, analyzed their performance and showed that they all failed below the same SNR. Our analysis revealed that the algorithm which first computed the SVD and then exploited the Kronecker structure performed the best.

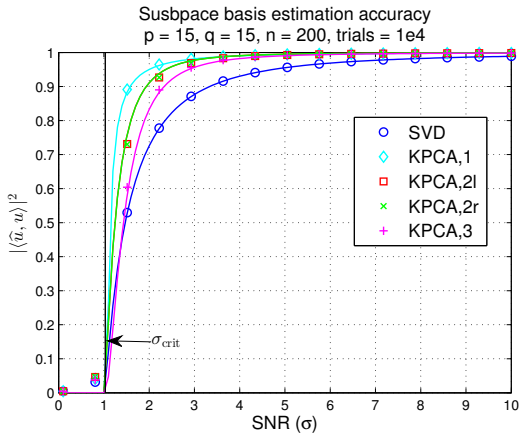


Fig. 1: Plot of inner-product-squared accuracy, $|\langle \widehat{u}, u\rangle|^2$, for a system with $q = 30, p = 30, n = 200$. The discrete points represent empirically simulated values, over 10^4 trials. The theoretical predictions are plotted using (14). For a given SNR σ_0, the accuracy of $\widehat{u}_{\text{kpca},1}$ is the highest, followed by $\widehat{u}_{\text{kpca},2\ell}$ and $\widehat{u}_{\text{kpca},2r}$ which are equal, followed by $\widehat{u}_{\text{kpca},3}$. All algorithms which exploit the Kronecker structure of u outperform the SVD in terms of $|\langle \widehat{u}, u\rangle|^2$ accuracy. $\sigma_{\text{crit}} = c_{\text{svd}}^{1/4}$ is the phase transition boundary indicated by (15).

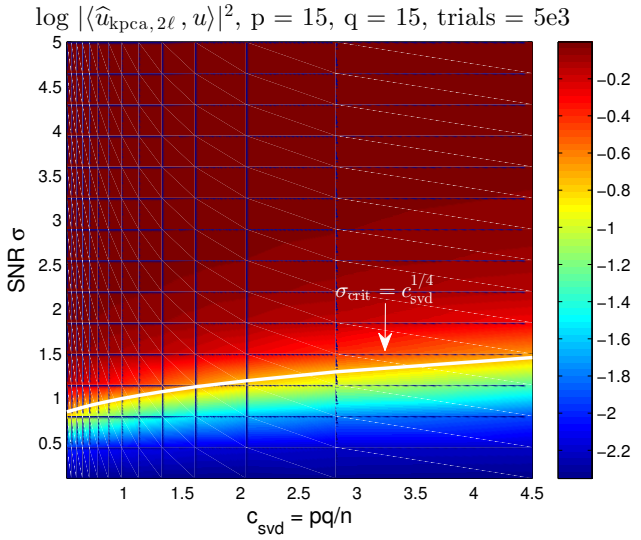


Fig. 2: Plot of the estimation accuracy $|\langle \widehat{u}, u\rangle|^2$ versus aspect ratio $c_{\text{svd}} = pq/n$ and SNR σ_0, for the system in Fig. 1. The red region is the regime where the $\widehat{u}$ is correlated with u, and the blue region is where the estimate is orthogonal to u, and the estimation performance is equivalent to choosing $\widehat{u}$ to be a random unit vector on the m dimensional sphere. We see a sharp phase transition boundary between the two regions. The phase transition boundary is predicted by (14) and (15) and accurately predicts the empirically observed boundary.

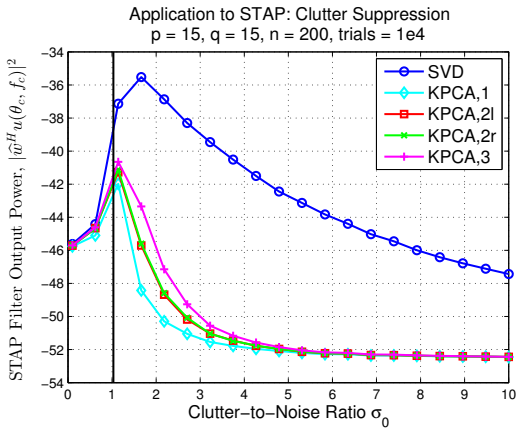
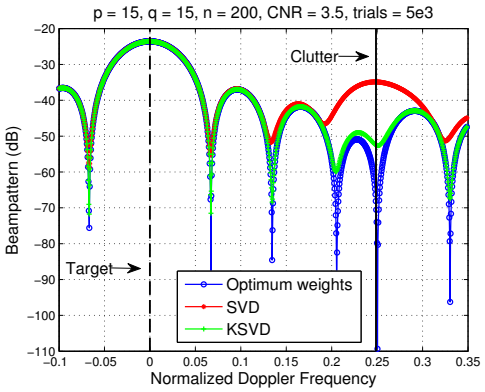
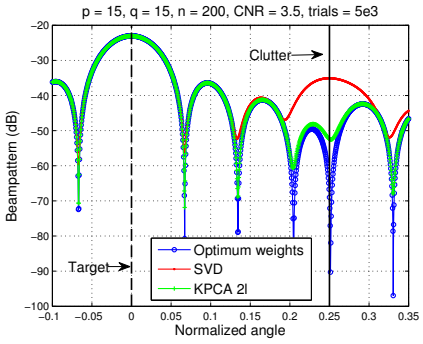


Fig. 3: Plot of the STAP filter output power at the true clutter location and Doppler frequency, i.e. $|\widehat{w}^H u(\theta_c, f_c)|^2$ in dB versus CNR σ_0, for the system described in Section 6. We see that the output power of the KPCA based algorithms is lesser than that of the SVD based filter estimate, indicating greater clutter suppression by about 10 dB in the low CNR regime.



(a) Beampattern (dB) vs normalized Doppler frequency at the true clutter location.



(b) Beampattern (dB) vs normalized array angle at the true Doppler frequency.

Fig. 4: Beampattern (dB) of the array, $|\langle \widehat{w}, u(\theta, f)\rangle|^2$ for varying θ at f_c and varying f at θ_c. In both plots, we see that the SVD and KSVD beampatterns both peak at the same level at the target location. However, at the clutter location, the notch due to the KPCA based filter leads to greater clutter suppression, as seen from Figure 3.

8. REFERENCES

[1] T. Tsiligkaridis and A. O. Hero, "Covariance estimation in high dimensions via kronecker product expansions," *Signal Processing, IEEE Transactions on*, vol. 61, no. 21, pp. 5347–5360, Nov 2013.

[2] C. F. Van Loan and N. Pitsianis, "Approximation with kronecker products," in *Linear Algebra for Large Scale and Real-Time Applications*, Marc S. Moonen, Gene H. Golub, and Bart L.R. De Moor, Eds., vol. 232 of *NATO ASI Series*, pp. 293–314. Springer Netherlands, 1993.

[3] Alan J. Laub, *Matrix Analysis For Scientists And Engineers*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2004.

[4] J. R. Guerci, *Space-time Adaptive Processing for Radar.*, Artech House, 2003.

[5] W. L. Melvin, "A STAP overview," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 19, no. 1, pp. 19–35, 2004.

[6] L. E. Brennan and I. S. Reed, "Theory of adaptive radar," *Aerospace and Electronic Systems, IEEE Transactions on*, , no. 2, pp. 237–252, 1973.

[7] Alexander Haimovich, "The Eigencanceler: Adaptive Radar by Eigenanalysis Methods," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 32, no. 2, pp. 532–542, 1996.

[8] F. Benaych-Georges and R. R. Nadakuditi, "The singular values and vectors of low rank perturbations of large rectangular random matrices," *Journal of Multivariate Analysis*, vol. 111, pp. 120–135, 2012.