Slow Radar Target Detection in Heterogeneous Clutter using Thinned STAP

Xiangrong Wang[1*], Elias Aboutanios[1], Moeness G. Amin[2]
[1]School of Electrical Engineering, University of New South Wales, Sydney, Australia 2052.
[2]Center for Advanced Communications, Villanova University, Villanova, PA 19085, USA.
[*]xiangrong.wang10@gmail.com.

Abstract

We address the problem of slow target detection in heterogeneous clutter through dimensionality reduction. Traditional approaches of implementing the space-time adaptive processing (STAP) require a large number of training data to estimate the clutter covariance matrix. In order to address the issue of limited training data especially in the heterogeneous scenarios, we propose a novel thinned STAP through selecting an optimum subset of antenna-pulse pairs that achieves the maximum output signal-to-clutter-plus-noise ratio (SCNR). The proposed strategy utilises a new parameter, named spatial spectrum correlation coefficient (S^2C^2), to analytically characterize the effect of space-time configuration on STAP performance and reduce the dimensionality of traditional STAP. Two algorithms are proposed to solve the antenna-pulse selection problem. The effectiveness of the proposed strategy is confirmed by extensive simulation results, especially by utilising the MCARM data set.

Index Terms

Spatial spectrum correlation coefficient, Heterogeneous clutter, MCARM, Combinatorial optimization, Antenna-pulse Selection

I. Introduction

Space-time adaptive processing (STAP) involves combining signals linearly from an array of antennas and multiple pulses to improve airborne radar detection performance in severe clutter and jamming environments [1]–[3]. The optimum STAP processor employs the clutter-plus-noise covariance matrix (CCM) to whiten the received data prior to the application of a matched-filter detector. As the CCM is usually unknown, practical STAP approaches such as the sample matrix inversion (SMI) method obtain the maximum likelihood estimate of the CCM from secondary cells. The number of independent and identically distributed (IID) training data, required by the SMI to ensure an average signal-to-clutter-plus-noise ratio (SCNR) loss within 3dB of the optimum processor, is twice the number of degrees of freedom (DoFs) of the detector. This is typically on the order of several hundreds for common STAP applications and can far exceed the available data measurements [4]. For non-homogeneous environment, the SMI may incur a significant performance degradation due to lack of sufficient IID training data [5].

In addition to clutter heterogeneity and limited sample support, practical implementations of STAP continue to face other challenges including high computational cost in real-world scenarios. Fortunately, most interference suppression problems in airborne radar systems are rank deficient in nature, that is, they require fewer adaptive DoFs compared to those offered by the array. By appropriate design, the redundant DoFs can be discarded and partial

STAP can be applied [6]. For instance, preprocessing can be used to select statistically representative training data in order to mitigate clutter heterogeneity [7]–[11]. Knowledge-aided (KA) STAP incorporates *a priori* knowledge into the estimation process to accelerate the convergence of the CCM [12]–[15]. Test data only algorithms such as the D^3 algorithm [16] and the MLED [17] do away with the need for training data. An image processing-based STAP technique was proposed in [18]. Principle Component Analysis (PCA) projects the data onto a lower dimensional subspace [19], [20] to reduce the high sample support. Some recent work has considered utilizing sparse recovery (SR) techniques of the clutter spectrum in the angle-Doppler plane for moving target indication [21], [22]. However, these SR approaches are computationally more expensive than conventional STAP because they require a full-dimensional matrix inversion and the recovery procedure calls for additional computations. The work in [23], [24] assumes the sparse property of the STAP filter weights instead of the clutter spectrum and utilized the l_1-norm to promote sparsity.

In this paper, we propose a novel thinned STAP approach consisting of an antenna-pulse selection strategy where we select a subset of optimal antenna-pulse pairs in each range gate prior to applying STAP and target detection. Traditional selection strategies viewed the antenna array separately from the pulse train and examined the selection on the array only in airborne radar to reject ground clutter [25], [26]. Whereas in this paper we view the interleaved antenna-pulse data as a virtual large array where each sensor is specified by the antenna number and pulse number. We carry out the "thinning" in both space and time through joint antenna-pulse selection to obtain a sub-configuration comprising the optimal sensor elements and consequently the optimal subset of antenna-pulse pairs. Compared to our work in [27], which implemented the antenna-pulse selection in terms of maximizing the separation between the target and the closest clutter Fourier basis, the contributions of this paper are: (1) We reformulate the parameter, referred to as spatial spectrum correlation coefficient (S^2C^2), to characterize the separation between the target and the clutter subspace and use it for analytically expressing the output SCNR. (2) We formulate and perform the antenna-pulse selection such that the resulting space-time configuration maximizes the output SCNR; (3) We propose two effective algorithms to implement the selection process; (4) We investigate the selection problem based on the eigenbasis, which exhibits slightly better performance than the Fourier basis based selection; (5) In addition to extensive simulations, the performance of the proposed antenna-pulse selection strategy is evaluated using monostatic Multi-Channel Airborne Radar Measurement (MCARM) data.

The paper is organized as follows: In section II, we review the clutter model and fundamental detection theory. In section IV, we introduce the parameter S^2C^2 and propose two approaches to solve the antenna-pulse selection problem. We also investigate the selection problem based on eigenbasis at the end of section IV. Section V gives simulation and experimental results that validate the performance of the proposed strategy. Finally, some conclusions are drawn in section VI.

II. Signal Model

The geometry of a sidelooking airborne array radar of N uniformly spaced antennas with inter-element spacing d in a Cartesian Coordinate system is shown in Fig. 3.2 in [2]. Let $\mathcal{P}$ be a scatterer patch on the ground having elevation angle θ and azimuth angle ϕ with respect to the array center. Consider an antenna with a distance $nd, n = 0, \cdots, N-1$ relative to the array origin. The signal received by the antenna from the stationary scatterer patch $\mathcal{P}$ is phase shifted relative to the origin by,

$$n2\pi f_s = n\frac{2\pi}{\lambda}d\cos\phi\cos\theta, \tag{1}$$

where $f_s = (d/\lambda)\cos\phi\cos\theta \in [-1/2, 1/2]$ is the normalized spatial frequency under the condition $d = \lambda/2$ and λ is the wavelength. The spatial steering vector can be expressed as,

$$\mathbf{a}_s = [1, e^{j2\pi f_s}, \cdots, e^{j(N-1)2\pi f_s}]^T. \tag{2}$$

In pulse Doppler radar, the Doppler frequency is measured by phase comparison between echo signals due to a transmitted coherent pulse train with pulse repetition interval (PRI) $\tilde{T}$. The unit phase shift incurred by the platform moving with the velocity v_p is given by

$$2\pi f_d = 2\pi\tilde{T}\frac{2v_p}{\lambda}\cos\phi\cos\theta. \tag{3}$$

Here we denote $f_d = (2v_p\tilde{T}/\lambda)\cos\phi\cos\theta$ as the normalized Doppler frequency. Then the temporal steering vector with M coherent pulses can be written as,

$$\mathbf{a}_t = [1, e^{j2\pi f_d}, \cdots, e^{j(M-1)2\pi f_d}]^T. \tag{4}$$

Thus, the interleaved space-time steering vector, $\mathbf{a}(\theta, \phi) \in \mathbb{C}^{NM\times 1}$, is

$$\mathbf{a}(\theta, \phi) = \mathbf{a}_s \otimes \mathbf{a}_t, \tag{5}$$

where $\otimes$ denotes Kronecker product. Consider a single range increment, specified by the elevation angle θ, the total clutter echo is an integral over the various contributions from all ground scatterers in azimuth,

$$\mathbf{c}(\theta) = \int_{\phi=0}^{2\pi} \mathcal{A}D(\phi, \theta)G(\phi, \theta)\mathbf{a}(\theta, \phi)d\phi, \tag{6}$$

where $\mathcal{A}$ stands for ground reflectivity that is assumed to be a circular complex Gaussian variable, $D(\phi, \theta)$ and $G(\phi, \theta)$ are the receive and transmit directivity patterns respectively. The trajectory of clutter spectrum in the angle-Doppler (f_s-f_d) plane is very important for extracting the information on clutter subspace. For a sidelooking array, the clutter trajectory is

$$f_d = \kappa f_s, \tag{7}$$

which is a straight line in the f_d-f_s plane with a slope $\kappa = 2v_p\tilde{T}/d$.

The detection problem is usually treated as a hypothesis test to determine the presence of a target. The model of the received signal, $\mathbf{x} \in \mathbb{C}^{NM}$, in a single range increment is,

$$H_0 : \mathbf{x} = \mathbf{c} + \mathbf{n}, \tag{8}$$

where H_0 denotes the null hypothesis. For simplicity of notation, we put $\mathbf{c}$ instead of $\mathbf{c}(\theta)$ and $\mathbf{n}$ denotes the Gaussian white noise with power σ_n^2. The alternative hypothesis is given by,

$$H_1 : \mathbf{x} = \alpha\mathbf{t} + \mathbf{c} + \mathbf{n}, \tag{9}$$

where α denotes the complex amplitude of target signal. The target space-time steering vector $\mathbf{t} \in \mathbb{C}^{NM\times 1}$ can be similarly constructed according to Eqs. (2), (4) and Eq. (5) with $f_s = (d/\lambda)\cos\phi_t\cos\theta_t$ and $f_d = 2v_t\tilde{T}/\lambda$ given a target in elevation θ_t and azimuth ϕ_t with radial velocity v_t relative to the platform.

We assume that all components of $\mathbf{x}$ are mutually independent. The CCM, $\mathbf{Q}$, is then a sum of clutter and noise

covariance matrices as follows,

$$\mathbf{Q} = E\{\mathbf{x}\mathbf{x}^H\} = \sigma_n^2 \mathbf{I}_{NM} + \mathbf{Q}_c, \tag{10}$$

where $\mathbf{Q}_c$ stands for the clutter covariance matrix and is usually rank-deficient. For a sidelooking linear array, Brennan's rule gives the clutter rank N_e to be, [1], [28],

$$N_e = \text{int}\{N + \kappa(M-1)\}, \tag{11}$$

where int$\{\}$ denotes the next integer number. Suppose the clutter rank is N_e, then

$$\mathbf{Q}_c = \sum_{i=1}^{N_e} \sigma_i^2 \mathbf{e}_i \mathbf{e}_i^H = \sum_{j=1}^{N_e} P_j \mathbf{v}_j \mathbf{v}_j^H, \tag{12}$$

where $\mathbf{e}_i$ and σ_i^2 are the ith eigenvector and its corresponding eigenvalue of clutter covariance matrix $\mathbf{Q}_c$ respectively. We call $\mathbf{e}_i, i = 1, \cdots, N_e$ as clutter eigenbasis in the following. The clutter subspace is also spanned by N_e Fourier basis vectors $\mathbf{v}_j, j = 1, \cdots, N_e$ with power coefficients P_j. The definition of Fourier basis $\mathbf{v}_j$ is same with that of the interleaved space-time steering vector $\mathbf{a}$ in Eq. (5) with its corresponding spatial and Doppler frequencies. Since the two sets of N_e basis vectors span the same clutter subspace, i.e. span$(\mathbf{e}_i, i = 1, \cdots, N_e) =$ span$(\mathbf{v}_j, j = 1, \cdots, N_e)$, each Fourier basis vector can be expressed as a linear combination of eigenbasis, i.e.,

$$\mathbf{v}_j = \sum_{i=1}^{N_e} \mu_i^j \mathbf{e}_i. \tag{13}$$

As the clairvoyant CCM is not available in practice, we adopt the Adaptive Matched Filter (AMF) detector in this work [29], which is formulated as

$$\frac{|\mathbf{v}^H \hat{\mathbf{Q}}^{-1} \mathbf{x}|^2}{\mathbf{v}^H \hat{\mathbf{Q}}^{-1} \mathbf{v}} \underset{H_1}{\overset{H_0}{\lessgtr}} \mathcal{T}, \tag{14}$$

where $\mathcal{T}$ is the threshold value and $\mathbf{v}$ is the scanning steering vector over the angle-Doppler plane. The CCM $\mathbf{Q}$ is estimated by $\hat{\mathbf{Q}} = \frac{1}{L}\sum_{l=1}^{L} \mathbf{x}(l)\mathbf{x}^H(l)$ given L homogeneous training data under the hypothesis H_0. Clearly, the maximum value of Eq. (14) occurs when $\mathbf{v} = \mathbf{t}$ in the cell under test (CUT).

III. Generalization of Spatial Spectral Correlation Coefficient

The spatial spectrum correlation coefficient (S^2C^2) was derived in a similar way to the single interference case in [27] and is therefore not suitable for the multi-interference case. In this section, we generalise the definition of S^2C^2 using the concept of the clutter subspace to effectively deal with multiple interferences [30].

A. Matrix-Vector Expression

Proceeding from Eq. (12), we arrange the N_e Fourier basis vectors into a matrix $\mathbf{V}_c \in \mathbb{C}^{MN \times N_e}$,

$$\mathbf{V}_c = [\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_{N_e}]. \tag{15}$$

Defining $\mathbf{P} = \text{diag}[P_1, \cdots, P_{N_e}]$, we rewrite the CCM in Eq. (10) as,

$$\mathbf{Q} = \sigma_n^2 \mathbf{I}_{NM} + \mathbf{V}_c \mathbf{P} \mathbf{V}_c^H. \tag{16}$$

Applying the Woodbury matrix identity to the inverse CCM, $\mathbf{Q}^{-1}$, yields,

$$\mathbf{Q}^{-1} = \frac{1}{\sigma_n^2}(\mathbf{I}_{NM} - \mathbf{V}_c(\sigma_n^2\mathbf{P}^{-1} + \mathbf{V}_c^H\mathbf{V}_c)^{-1}\mathbf{V}_c^H). \tag{17}$$

We assume the clutter is much stronger than white noise, i.e. $P_1 > \cdots > P_{N_c} \gg \sigma_n^2$, then Eq. (17) can be further simplified as

$$\mathbf{Q}^{-1} \approx \frac{1}{\sigma_n^2}(\mathbf{I}_{NM} - \mathbf{V}_c(\mathbf{V}_c^H\mathbf{V}_c)^{-1}\mathbf{V}_c^H). \tag{18}$$

From the above equation it is clear that when the CNR is high, $\mathbf{Q}^{-1}$ approximates the clutter nullspace. Accordingly, the optimum STAP weight vector, [2], [29],

$$\mathbf{w}_{\text{opt}} = \eta\mathbf{Q}^{-1}\mathbf{t} \approx \frac{\eta}{\sigma_n^2}(\mathbf{I}_{NM} - \mathbf{V}_c(\mathbf{V}_c^H\mathbf{V}_c)^{-1}\mathbf{V}_c^H)\mathbf{t}, \tag{19}$$

becomes the interference eigencanceler proposed in [19]. Here $\eta = (\mathbf{t}\mathbf{Q}^{-1}\mathbf{t})^{-1/2}$ does not affect the output SCNR.

The steering vector of the target signal $\mathbf{t}$ can be decomposed into two orthogonal components: the clutter subspace component $\mathbf{t}_c$ and the nullspace component $\mathbf{t}_\perp$, i.e., $\mathbf{t} = \mathbf{t}_c + \mathbf{t}_\perp$ with $\mathbf{t}_c = (\mathbf{V}_c(\mathbf{V}_c^H\mathbf{V}_c)^{-1}\mathbf{V}_c^H)\mathbf{t}$ and $\mathbf{t}_\perp = (\mathbf{I}_{NM} - \mathbf{V}_c(\mathbf{V}_c^H\mathbf{V}_c)^{-1}\mathbf{V}_c^H)\mathbf{t}$ respectively. The STAP weight vector $\mathbf{w}_{\text{opt}}$, is along the clutter nullspace direction $\mathbf{t}_\perp$. We define the S^2C^2 as the absolute value of the cosine of the angle between the target signal $\mathbf{t}$ and clutter subspace component $\mathbf{t}_c$, i.e.,

$$|\alpha| = |\cos\vartheta| = \left|\frac{\mathbf{t}^H\mathbf{t}_c}{\|\mathbf{t}\|_2\|\mathbf{t}_c\|_2}\right|. \tag{20}$$

Here the length of $\mathbf{t}$ is $\|\mathbf{t}\|_2 = \sqrt{MN}$ under the assumption of isotropic antennas. As the SCNR_{out} is directly related to the squared value of the S^2C^2, we substitute the expression of $\mathbf{t}_c$ into Eq. (20) and take the squared value yields,

$$\begin{aligned} |\alpha|^2 &= \frac{\left|\mathbf{t}^H\mathbf{V}_c(\mathbf{V}_c^H\mathbf{V}_c)^{-1}\mathbf{V}_c^H\mathbf{t}\right|^2}{MN\|\mathbf{V}_c(\mathbf{V}_c^H\mathbf{V}_c)^{-1}\mathbf{V}_c^H\mathbf{t}\|_2^2}, \\ &= \frac{1}{MN}\mathbf{t}^H\mathbf{V}_c(\mathbf{V}_c^H\mathbf{V}_c)^{-1}\mathbf{V}_c^H\mathbf{t}. \end{aligned} \tag{21}$$

Finally, combining Eqs. (18) and (21), the output SCNR (SCNR_{out}) becomes

$$\begin{aligned} \text{SCNR}_{\text{out}} &= \sigma_t^2\mathbf{t}^H\mathbf{Q}^{-1}\mathbf{t} \\ &\approx \frac{\sigma_t^2}{\sigma_n^2}\mathbf{t}^H(\mathbf{I}_{NM} - \mathbf{V}_c(\mathbf{V}_c^H\mathbf{V}_c)^{-1}\mathbf{V}_c^H)\mathbf{t} \\ &\approx \text{SNR}\cdot MN(1-|\alpha|^2), \end{aligned} \tag{22}$$

where σ_t^2 denotes the power of the target signal and $\text{SNR} = \sigma_t^2/\sigma_n^2$. Eq. (22) shows that the SCNR_{out} depends on two factors: the number of available DoFs (which is MN for the full configuration) and the squared S^2C^2 value $|\alpha|^2$. When the number of available DoFs is fixed, the performance can be improved by changing the space-time configuration to reduce the S^2C^2 value. Thus the S^2C^2 characterizes the effect of the space-time geometry on the adaptive filtering performance and, in this respect, is an effective metric for optimum antenna-pulse selection.

B. Determinant Expression

The matrix-vector expression of the S^2C^2 given in Eq. (21) is not a convenient form for antenna-pulse selection. Below we derive a compact formula of the S^2C^2 based on Eq. (21) in terms of matrix determinants.

Let the clutter cross-correlation matrix $\mathbf{D}_c \in \mathbb{C}^{N_e \times N_e}$ be

$$\begin{aligned} \mathbf{D}_c &= \mathbf{V}_c^H \mathbf{V}_c \\ &= \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1N_e} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2N_e} \\ \cdots & \cdots & \cdots & \cdots \\ \rho_{N_e1} & \rho_{N_e2} & \cdots & \rho_{N_eN_e} \end{bmatrix}, \end{aligned} \tag{23}$$

where the entry $\rho_{ij} = \mathbf{v}_i^H \mathbf{v}_j$ for $i, j = 1, ..., N_e$. Note that the estimated Fourier basis vectors ($\hat{\mathbf{v}}_j$ in Eq. (39)) are not orthogonal in practical applications, thus we do not simplify $\mathbf{D}_c$ as $\mathbf{I}_{N_e}$ for generality. Define $\mathbf{V}_t = [\mathbf{t}, \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_{N_e}]$, the target plus clutter cross-correlation matrix $\mathbf{D}_t \in \mathbb{C}^{(N_e+1)\times(N_e+1)}$ is

$$\begin{aligned} \mathbf{D}_t &= \mathbf{V}_t \mathbf{V}_t^H \\ &= \begin{bmatrix} \rho_{tt} & \rho_{t1} & \cdots & \rho_{tN_e} \\ \rho_{1t} & \rho_{11} & \cdots & \rho_{1N_e} \\ \cdots & \cdots & \cdots & \cdots \\ \rho_{N_et} & \rho_{N_e1} & \cdots & \rho_{N_eN_e} \end{bmatrix}, \\ &= \begin{bmatrix} MN & \mathbf{t}^H \mathbf{V}_c \\ \mathbf{V}_c^H \mathbf{t} & \mathbf{D}_c \end{bmatrix}, \end{aligned} \tag{24}$$

where the entry $\rho_{tj} = \rho_{jt}^* = \mathbf{t}^H \mathbf{v}_j$ for $j = 1, ..., N_e$. Utilizing the determinant formula of a block matrix in Eq. (24) yields,

$$|\mathbf{D}_t| = |\mathbf{D}_c|(MN - \mathbf{t}^H \mathbf{V}_c \mathbf{D}_c^{-1} \mathbf{V}_c^H \mathbf{t}). \tag{25}$$

Thus,

$$\mathbf{t}^H \mathbf{V}_c \mathbf{D}_c^{-1} \mathbf{V}_c^H \mathbf{t} = MN - \frac{|\mathbf{D}_t|}{|\mathbf{D}_c|}. \tag{26}$$

Substituting Eqs. (23) and (26) into Eq. (21), the expression of the $\mathrm{S}^2\mathrm{C}^2$ can be rewritten in terms of the ratio of two matrix determinants,

$$|\alpha|^2 = \frac{1}{MN} \mathbf{t}^H \mathbf{V}_c \mathbf{D}_c^{-1} \mathbf{V}_c^H \mathbf{t} = 1 - \frac{|\mathbf{D}_t|}{MN|\mathbf{D}_c|}. \tag{27}$$

In the case of single interference as in [31], i.e. when $N_e = 1$, the two cross-correlation matrices in Eqs. (23) and (24) reduce to $\mathbf{D}_c = \rho_{11} = MN$ and

$$\mathbf{D}_t = \begin{bmatrix} MN & \rho_{t1} \\ \rho_{1t} & MN \end{bmatrix}, \tag{28}$$

respectively. Substituting $|\mathbf{D}_c| = MN$ and $|\mathbf{D}_t| = M^2N^2 - |\rho_{t1}|^2$ into Eq. (27) yields $|\alpha|^2 = |\rho_{t1}|^2/(M^2N^2)$. Thus this generalized formula is in accordance with the formula of the single interference case given by Eq. (11) in [31].

Now, we proceed to substitute Eq. (27) into Eq. (22),

$$\mathrm{SCNR}_{\mathrm{out}} \approx \mathrm{SNR} \cdot MN(1 - |\alpha|^2) \approx \mathrm{SNR} \cdot \frac{|\mathbf{D}_t|}{|\mathbf{D}_c|}. \tag{29}$$

Note that Eq. (29) is equivalent to Eq. (22), but there is no explicit linear dependence between the performance

and the number of selected antenna-pulse pairs in Eq. (29). This further demonstrates the non-linear relationship between the SCNR_{out} and the number of DoFs. In section IV, we implement antenna-pulse selection in terms of maximizing the SCNR_{out} in Eq. (29) directly.

IV. Antenna-Pulse Selection for Clutter Suppression

We propose two selection algorithms in terms of maximizing the output SCNR in this section. Subsequently we present two ways to estimate the clutter subspace.

A. D.C. Programming

Define the binary selection vector $\mathbf{z} \in \{0,1\}^{MN}$ with "one" meaning the corresponding antenna-pulse pair is selected and "zero" meaning discarded. Then, the two cross-correlation matrices of the selected subarray can be expressed as

$$\mathbf{D}_c(\mathbf{z}) = \mathbf{V}_c^H \text{diag}(\mathbf{z}) \mathbf{V}_c,\ \mathbf{D}_t(\mathbf{z}) = \mathbf{V}_t^H \text{diag}(\mathbf{z}) \mathbf{V}_t, \tag{30}$$

with the two matrices $\mathbf{V}_c$ and $\mathbf{V}_t$ defined in Eqs. (15) and (24). Note that $\mathbf{D}_c(\mathbf{z})$ is not an identity matrix after implementing selection although the N_e clutter basis vectors of the full configuration are mutually orthogonal. This is the second reason why we could not simplify $\mathbf{D}_c$ as $\mathbf{I}_{N_e}$ in Eq. (23). Correspondingly, the SCNR_{out} of the selected sub-configuration can be written as

$$\text{SCNR}_{\text{out}} = \text{SNR} \cdot \frac{|\mathbf{D}_t(\mathbf{z})|}{|\mathbf{D}_c(\mathbf{z})|} = \text{SNR} \cdot \frac{|\mathbf{V}_t^H \text{diag}(\mathbf{z}) \mathbf{V}_t|}{|\mathbf{V}_c^H \text{diag}(\mathbf{z}) \mathbf{V}_c|}. \tag{31}$$

The antenna-pulse selection problem in terms of maximizing the SCNR_{out} is formulated as,

$$\begin{aligned} \min_{\mathbf{z}} \quad & \frac{|\mathbf{D}_c(\mathbf{z})|}{|\mathbf{D}_t(\mathbf{z})|}, \\ \text{s.t.} \quad & \mathbf{z} \in \{0,1\}^{MN}. \end{aligned} \tag{32}$$

Note that the objective function in Eq. (32) is equivalent to $\max_{\mathbf{z}} \text{SCNR}_{\text{out}}$. It is clear that the global minimizer is a vector of all ones, when no restriction is placed on the number of selected entries. As it is critical in STAP that the number of DoFs be matched to the available training data [32], we constrain the number of selected antenna-pulse pairs to be K, given $L = 2K$ training data. The optimization problem becomes,

$$\begin{aligned} \min_{\mathbf{z}} \quad & \frac{|\mathbf{D}_c(\mathbf{z})|}{|\mathbf{D}_t(\mathbf{z})|}, \\ \text{s.t.} \quad & \mathbf{z} \in \{0,1\}^{MN}, \\ & \mathbf{1}^T \mathbf{z} = K. \end{aligned} \tag{33}$$

Define the feasible set $\mathcal{S} = \{\mathbf{z} : \mathbf{z} \in \{0,1\}^{MN}\}$, which comprises the extreme points of the polytope $\mathcal{D} = \{\mathbf{z} : 0 \preceq \mathbf{z} \preceq 1\}$, and relax the constraint $\mathbf{z} \in \mathcal{S}$ to $\mathbf{z} \in \mathcal{D}$. Since both $\mathbf{D}_c(\mathbf{z})$ and $\mathbf{D}_t(\mathbf{z})$ are positive definite and the logarithm function is monotonically increasing [33], Eq. (33) is relaxed into the following problem,

$$\begin{aligned} \min_{\mathbf{z}} \quad & \log(|\mathbf{D}_c(\mathbf{z})|) - \log(|\mathbf{D}_t(\mathbf{z})|), \\ \text{s.t.} \quad & \mathbf{1}^T \mathbf{z} = K, \\ & \mathbf{z} \in \mathcal{D}. \end{aligned} \tag{34}$$

The objective function of Eq. (34) is a difference of two concave functions and belongs to D.C. Programming [34]–[36]. A convex-concave procedure (CCP) is adopted to solve it [37], which is proven to be able to converge to a stationary point in [38], [39] with a feasible initial point. Although the stationary point may not be necessarily the local minimum, it was proven to converge to the global optimum under certain conditions in [36]. Note that the DCA (d.c. algorithm) in [34], [35] is reduced to the CCP for differentiable optimization problems. In the following, we give a detailed implementation of the CCP.

The concave function $f(\mathbf{z}) = \log(|\mathbf{D}_c(\mathbf{z})|)$ is approximated iteratively by its first-order Taylor decomposition in the $(k+1)$th iteration as

$$f(\mathbf{z}) \approx \hat{f}(\mathbf{z}) = f(\mathbf{z}^{(k)}) + \nabla f(\mathbf{z}^{(k)})^T(\mathbf{z} - \mathbf{z}^{(k)}). \tag{35}$$

The jth entry of the gradient $\nabla f(\mathbf{z}^{(k)})$ is

$$\nabla f_j(\mathbf{z}^{(k)}) = \text{tr}\left\{\mathbf{D}_c^{-1}(\mathbf{z}^{(k)})(\mathbf{v}_{c,j}\mathbf{v}_{c,j}^H)\right\}, \tag{36}$$

where the operator tr$\{\bullet\}$ takes the trace of the matrix $\bullet$ and $\mathbf{v}_{c,j} \in \mathbb{C}^{N_e \times 1}$ is the jth row vector of $\mathbf{V}_c$. Substituting Eqs. (35), (36) into Eq. (34) and dropping the constant terms, the $(k+1)$th iteration is formulated as,

$$\begin{aligned} \min_{\mathbf{z}} \quad & \nabla f(\mathbf{z}^{(k)})^T\mathbf{z} - \log(|\mathbf{D}_t(\mathbf{z})|), \\ \text{s.t.} \quad & \mathbf{1}^T\mathbf{z} = K, \\ & \mathbf{z} \in \mathcal{D}. \end{aligned} \tag{37}$$

According to [40], [41], the global optimum solution of a D.C. programming is on the edge of the polytope $\mathcal{D}$, which is sparse but not necessarily binary. Solving DC problems most often relies on branch and bound methods or cutting plane methods, whereas these global methods often prove slow in practice, requiring many partitions or cuts. Therefore, we are instead concerned with local heuristic that can find local optimum binary solutions rapidly.

Inspired by the effective Gaussian randomization approach in [42], we modify the randomization procedure to obtain a binary solution. We assume a random vector ξ with each component $\xi_i \sim \mathcal{N}(\hat{\mathbf{z}}_i, \varepsilon_i)$, where the mean $\hat{\mathbf{z}}$ is the optimum solution of Eq. (37) and ε_i denotes the variance of the ith entry ξ_i. After obtaining a random vector ξ by sampling $\mathcal{N}(\hat{\mathbf{z}}, \text{diag}(\varepsilon))$, we then set the first K largest entries to one and others to zero to generate a feasible point. Moreover, we repeat this random sampling multiple times and choose the one that yields the best objective. As shown in [42], randomization provides an effective approximation for a small number of randomizations. The same conclusion applies here. In order to preserve the one and zero entries of $\hat{\mathbf{z}}$ while varying the non-binary entries, we set $\varepsilon_i = \hat{\mathbf{z}}_i(1 - \hat{\mathbf{z}}_i)$ with the highest variance 0.25 for the most ambiguous entry 0.5. We point out that D.C. Programming can also be adapted to find the sparsest configuration with the required performance. This is achieved by swapping the output SCNR term in the objective function with the number constraint in the second line of Eq. (37).

B. Modified Correlation Measurement

Since the CCP requires solving a sequence of convex optimization problems with computational complexity of order $O((MN)^{3.5})$ for each iteration in the worst case scenario, D.C. Programming is computationally expensive although with theoretical convergence guarantee. We devise a simple alternative search method, named Modified Correlation Measurement (MCM) based on [31]. The MCM belongs to greedy search algorithm which is effective

TABLE I
MODIFIED CORRELATION MEASUREMENT METHOD

Step 1	Select all antenna-pulse pairs, i.e. $\mathbf{z} = \mathbf{1}$, iteration number $k = 1$, index set $\chi^{(1)} = [1, \cdots, MN]$;
Step 2	For $l = 1 : MN - k + 1$ Set $\tilde{\mathbf{z}} = \mathbf{z}$ and $\tilde{\mathbf{z}}(\chi^{(k)}(l)) = 0$, Calculate $r(l) = \frac{\lvert\mathbf{D}_c(\tilde{\mathbf{z}})\rvert}{\lvert\mathbf{D}_t(\tilde{\mathbf{z}})\rvert}$, End;
Step 3	Obtain $i = \arg\min_{l=1,\ldots,MN-k+1} r(l)$;
Step 4	Set $\mathbf{z}(\chi^{(k)}(i)) = 0$, and update the index set $\chi^{(k+1)} = \chi^{(k)} \setminus \chi^{(k)}(i) = \{n \in \chi^{(k)}, n \neq \chi^{(k)}(i)\}$;
Step 5	$k = k + 1$, If $k = MN - K$, terminate; otherwise go back to Step 2.

in solving combinatorial optimization problems although without complete mathematical proof up to now [43]. The detailed implementation procedure of the MCM is presented in Table I. Basically, all antenna pulse pairs are selected initially and a backward search is utilized to discard the antenna-pulse pair that gives the smallest objective value in Eq. (33) in each iteration until K antenna-pulse pairs remain.

In practice we may need to obtain the sparsest space-time configuration which satisfies the performance requirement. Since the output SCNR is monotonically increasing with the number of selected antenna-pulse pairs, the MCM can be adapted to achieve this aim. Instead of enforcing constraints on the number of selected antenna-pulse pairs, the termination condition in Step 5 can be the reduction of the output SCNR.

Now Let us calculate the computational complexity of the MCM method. As there are in total $MN - K$ steps in MCM and we need to calculate the determinant of a matrix with dimension $N_e \times N_e$ in each step, the computational complexity of the MCM is thus of the order $O(N_e^3 \times (MN - K))$. Taking the filter weight calculation into account, i.e. implementing STAP with K antenna-pulse pairs, the total computational load is of the order $O(N_e^3 \times MN + K^3)$, which is smaller than that of the traditional STAP of complexity $O(M^3N^3)$. The reason is that the effective clutter rank $N_e \ll MN$, such as $N_e = M + N$ for the sidelooking array. More importantly, the proposed selection strategy is most advantageous for detecting slow-moving targets with reduced training data, which means it is not necessary to implement selection for each angle-Doppler bin.

C. Clutter Subspace Estimation

The implementation of antenna-pulse selection can be directly performed by the element-wise product $\tilde{\mathbf{v}}_j = \mathbf{z} \odot \mathbf{v}_j, j = 1, \cdots, N_e$. Thus it is intuitive that the reduced dimensional clutter subspace can be characterized by $\tilde{\mathbf{v}}_j, i = 1, \cdots, N_e$ after selection, i.e. $\tilde{\mathbf{V}}_c = [\tilde{\mathbf{v}}_1, \cdots, \tilde{\mathbf{v}}_{N_e}]$ spans the reduced dimensional clutter subspace. Since the accuracy of Fourier analysis is restricted by its leakage effect, we prefer the eigenbasis to describe the clutter subspace. Substituting Eq. (13) into $\tilde{\mathbf{v}}_j$, we have that

$$\tilde{\mathbf{v}}_j = \sum_{i=1}^{N_e} \mu_i^j (\mathbf{z} \odot \mathbf{e}_i) = \sum_{i=1}^{N_e} \mu_i^j \tilde{\mathbf{e}}_i, j = 1, \cdots, N_e. \tag{38}$$

Therefore, the set of selected eigenbasis $\tilde{\mathbf{E}}_c = [\tilde{\mathbf{e}}_1, \cdots, \tilde{\mathbf{e}}_{N_e}]$ also spans the reduced dimensional clutter subspace, i.e. $\text{span}(\tilde{\mathbf{E}}_c) = \text{span}(\tilde{\mathbf{V}}_c)$. We can replace $\mathbf{V}_c = \mathbf{E}_c$ and set $\mathbf{V}_t = [\mathbf{E}_c, \mathbf{t}]$ to carry out the antenna-pulse selection.

Since the two sets of exact clutter basis $\mathbf{e}_i, i = 1, \cdots, N_e$ and $\mathbf{v}_j, j = 1, \cdots, N_e$ are not known as a prior, we need to estimate them in practice. The estimate of Fourier basis can be obtained by applying Fourier analysis to the received data in the CUT, i.e.,

$$\hat{\mathbf{v}} = \text{argmax}_{\mathbf{v}} |\mathbf{v}^H \mathbf{x}|. \tag{39}$$

The estimated power coefficient is then $\hat{P} = |\hat{\mathbf{v}}^H \mathbf{x}|^2$. The steering vector $\mathbf{v}$ is scanning the whole angle-Doppler plane, which is covered by MN resolution grids with N and M uniformly distributed spatial and Doppler normalized frequencies respectively. As shown in [44], for large enough clutter-to-noise ratio (CNR), the clutter subspace is well defined by Fourier basis vectors, and the clutter rank corresponds to the number of angle-Doppler grids where the clutter power is significant. The estimation of clutter eigenbasis requires training data to perform eigenvalue decomposition on the estimated covariance matrix $\hat{\mathbf{Q}}$.

We give a comparison between the eigenbasis and Fourier basis for antenna-pulse selection. Clearly, eigenvalue decomposition is computationally expensive. The set of clutter Fourier basis vectors is easy to compute and only the received data in the CUT is sufficient, whereas additional training data is required for computing the eigenbasis. As there is no accurate threshold to define the set of significant clutter Fourier basis [44], the set of eigenvectors achieves a better characterization of the clutter subspace and exhibits more robustness than the Fourier basis, thus the selected sub-configuration based on the eigenbasis would be expected to exhibit slightly better performance than that obtained using the Fourier basis, which can be observed from the simulation results in subsection V-B.

V. Simulation and Experimental Results

We now validate the proposed thinned STAP using simulated (subsections V-A to V-D) and experimental (subsection V-E) data. We utilize the MCM method to calculate optimum sub-configurations and the CNR is fixed at 20 dB throughout this section.

A. Validation of two proposed methods

In order to demonstrate the effectiveness of the proposed two antenna-pulse selection methods, we use a small radar system which permits us to enumerate all possible sub-configurations. Thus the radar has 4 antennas and 4 pulses and we select 10 antenna-pulse pairs out of 16. We take the sidelooking array as an example in the whole simulation, even though the proposed antenna-pulse selection strategy has no restriction on the array orientation. We inject a target at a broadside, that is $f_s = 0$. The Doppler frequency is swept over the range $[-0.5, 0.5]$. An optimum subset of 10 antenna-pulse pairs is selected through both proposed selection methods as well as enumeration. In Fig. 1, we plot the normalized distance D_e between the evaluated $\tilde{\text{SCNR}}_{\text{out}}$ and the true value SCNR_{out} as a function of f_d, i.e.,

$$D_e = \frac{\text{SCNR}_{\text{out}} - \tilde{\text{SCNR}}_{\text{out}}}{\text{SCNR}_{\text{out}}}, \tag{40}$$

where $\tilde{\text{SCNR}}_{\text{out}}$ is the output SCNR evaluated at the obtained solution and SCNR_{out} denotes the true optimum value though enumeration. We can see that both methods are capable of finding the global optimum configuration for slow-moving targets. In other scenarios the sub-optimum solutions can also produce acceptable performance through empirical simulations.

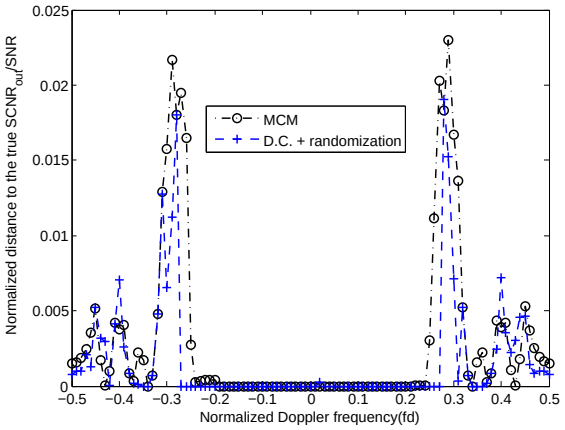

Fig. 1. Validation of the MCM, D.C. programming and randomization.

B. Redundancy of Full Configuration

A large array of uniformly spaced elements often exhibits significant redundancy, meaning that dropping the redundant elements hardly affects performance. To show this, we now use a radar platform with 11 antennas and 20 pulses and examine the relationship between the probability of detection and the number of selected antenna-pulse pairs K. Again, we run the simulation for both the Fourier and eigenbasis in Fig. 2. We consider two scenarios where the target normalized Doppler frequency is $f_d = 0.05$ and $f_d = 0.1$ respectively. The spatial frequency of the target is set as zero. The output SCNR value is set to -5dB for the full configuration and homogeneous clutter is assumed. For each K, the optimum configuration of K antenna-pulse pairs is calculated. We utilize $L = 2K$ IID training data to calculate the CCM in the outer plot and fix $L = 220$ in the inner plot. We can see that the optimum sub-configuration achieves the same detection performance as the full one with a much smaller number of DoFs from the outer plot, only about 80 antenna-pulse pairs when $f_d = 0.05$. The number increases to 120 when $f_d = 0.1$. The results demonstrate the significant redundancy of the full configuration, especially for slow-moving targets. When there is a fixed number of training data, such as $L = 220$ in the inner plot, the best performance is achieved in the neighbourhood of $K = L/2$. Interestingly, an optimum sub-configuration with 50 antenna-pulse pairs is sufficient when $f_d = 0.05$, whereas the detection probability is only 65% when $f_d = 0.1$. This further confirms the efficiency of the proposed thinned STAP for detecting slow-moving targets. Also observe that there is slight performance difference between the eigenbasis based selection and that of the Fourier basis, which justifies the use of the Fourier basis especially that it is computationally much simpler. The antenna-pulse selection is carried out on the Fourier basis in what follows.

C. Normalized SCNR Loss

We still assume homogeneous clutter returns in this simulation. The number of IID training data required by the SMI algorithm for the full configuration to achieve an SCNR loss within 3 dB of the optimum is $2NM = 440$. But we set the number of selected antenna-pulse pairs to $K = 100$, meaning that only $2K = 200$ training data snapshots are required for the sub-configurations. The performance of clutter suppression strategies in STAP is

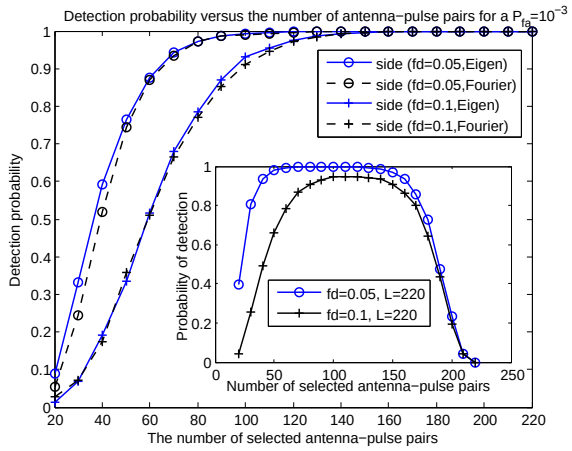

Fig. 2. The detection probability versus different numbers of antenna-pulse pairs. The number of training data keeps twice of that of antenna-pulse pairs, i.e. $L = 2K$ in the outer plot; whereas the number of training data is fixed as $L = 220$ in the inner. Each point is averaged over 10^5 Monte-Carlo runs.

commonly assessed using the normalized SCNR loss, [3],

$$L_s = \frac{|\mathbf{v}^H \hat{\mathbf{Q}}^{-1} \mathbf{v}|^2}{|\mathbf{v}^H \hat{\mathbf{Q}}^{-1} \mathbf{Q} \hat{\mathbf{Q}}^{-1} \mathbf{v}||\mathbf{v}^H \mathbf{Q}^{-1} \mathbf{v}|}, \tag{41}$$

where $\hat{\mathbf{Q}} = \frac{1}{L} \sum_{i=1}^{L} \mathbf{x}_i \mathbf{x}_i^H$ is the sample CCM and $\mathbf{v}$ is the scanning steering vector. Firstly, we examine the CCM convergence rate in the upper plot of Fig. 3. The normalized Doppler frequency of the target is set to $f_d = 0.05$, where the corresponding optimum sub-configuration is shown in Fig. 4. Clearly, the extreme antennas and pulses on the boundary are always included to preserve the maximum spatio-temporal aperture length, which directly determines the minimum detectable velocity. The number of homogeneous training data is changing from 100 for the sub-configuration (220 for the full configuration to avoid singularity) to 440 in steps of 10. We can see that the CCM of the full configuration requires at least 300 IID training data to converge, which is much slower than the sub-configuration. We then examine the relationship between L_s and the normalized Doppler frequency in the lower plot of Fig. 3. The normalized Doppler frequency of the target sweeps over the range $[-0.5, 0.1]$ and we calculate the optimum subset of 100 antenna-pulse pairs corresponding to each frequency. We compare the optimum configuration, which we denote by "sub$_{\text{opt}}$" using 200 training data, to the full configuration, denoted by "full" using 440 training data. The performances of the full and sub-configuration with 300 training data are also included for reference. We can observe the considerable performance degradation for the full configuration without sufficient training data.

D. Robustness against Heterogeneity

Finally we illustrate the improvement of the detection probability that the selection strategy provides in heterogeneous clutter. In the simulation, the probability of false alarm rate is set to $P_{\text{fa}} = 10^{-3}$. As we are primarily interested in slow moving targets, we assume the normalised target Doppler frequency to be uniformly distributed over the range $[-0.1, 0.1]$. We use the label "homo" to refer to curves for a homogeneous environment, leaving all others that show the performance under heterogeneity without a label. In the homogeneous case, we obtain the detection curves for the full configuration when the CCM is estimated from 440 IID training data snapshots. In the heterogeneous scenario, we simulate the clutter heterogeneity by inserting 60 high-amplitude, mainbeam

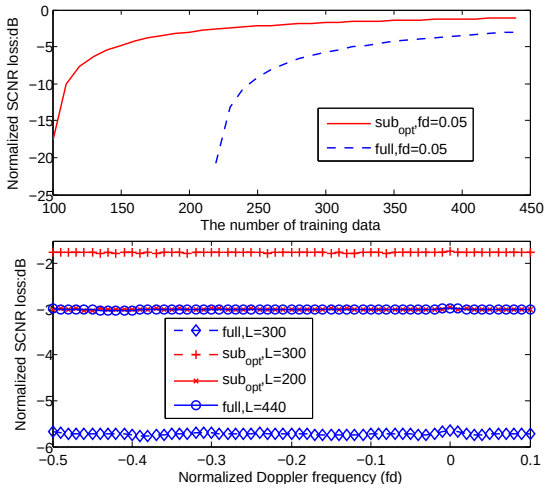

Fig. 3. Normalized SCNR loss versus the number of training data with $f_d = 0.05$ (upper plot) and versus normalised Doppler frequency with $L = 2K \setminus 300$ (lower plot), each point is averaged over 10^4 Monte-Carlo runs.

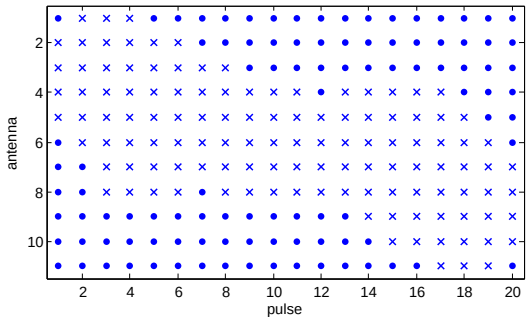

Fig. 4. Selected antenna-pulse pairs in the cases of $f_d = 0.05$. The filled dots stand for selected and the cross means discarded.

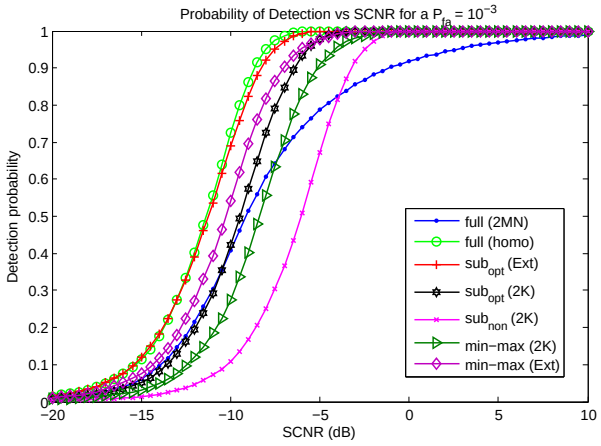

Fig. 5. Detection probability curve versus SCNR in both homogeneous and heterogeneous clutter, each point is averaged over 10^5 Monte-Carlo runs.

discrete targets into various range cells with normalized spatial and Doppler frequencies uniformly distributed within $[-0.5, 0.5]$, [7]. The required training data snapshots for both sub-configurations are selected using the generalized inner product (GIP)-based non-homogeneity detector (NHD) of [7]. Specifically, we sort the GIP values and retain the $2K = 200$ realizations corresponding to smallest GIP value. Moreover, the labels $2K$ and $2MN$ indicate the

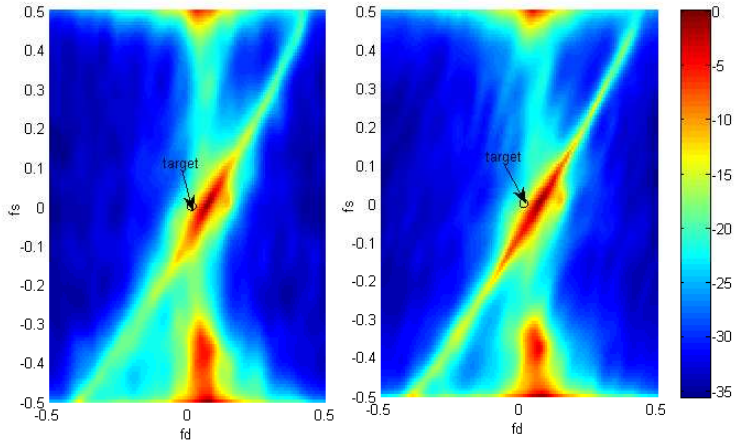

Fig. 6. Clutter spectrum in angle-Doppler plane for both full configuration (left) and the selected optimum sub-configuration (right).

number of snapshots used for the CCM estimation and the curve labelled "Ext" is the result of using all the snapshots that are returned by the NHD with GIP value below some threshold. In order to better exhibit the superiority of the proposed selection method, we compare its detection performance with the iterative min-max algorithm [27], denoted as "min-max". Note that the sub_{non} is the sub-configuration with the first 100 antenna-pulse pairs selected. We make the following observations:

- The sub-configurations combined with the NHD retain practically the same detection performance as the homogeneous case. We do not plot the curves for the sub-configurations in the homogeneous case for simplicity.
- The proposed subspace based selection method outperforms the iterative min-max algorithm, as the generalized S^2C^2 directly relates the output SCNR with the selected antenna-pulse pairs;
- The optimum sub-configuration exhibits better performance than the non-optimum and min-max configurations in both homogeneous and heterogeneous clutter;
- In heterogeneous clutter, the optimum sub-configuration combined with the NHD outperforms the full array, as it requires fewer IID training data than the full one.
- Interestingly, we find that the optimum sub-configuration with extended IID training data exhibits similar performance with the full configuration in the homogeneous clutter, which further illustrates the redundancy of the full configuration in the scenarios where the target is close to the clutter trajectory.

E. Experimental Results

The experimental radar data we used is the MCARM database [45]. The radar platform was travelling at around 100ms^{-1} over an area that includes a variety of terrain types as well as interfering discretes. The radar was comprised of a rectangular array of two rows in elevation and 11 sensors in azimuth and collected a coherent pulse interval of length 128 pulses. The pulse repetition frequency was 1984 Hz. The examples presented here are based on the acquisition 575 from flight 5. Since the number of available range cells is only 630, we employ the upper 11 uniformly spaced antennas and the first 20 pulses which satisfies the sample size requirement for the CCM estimation. The clutter spectrum in the angle-Doppler plane is estimated in the left plot of Fig. 6. The clutter ridge is shifted right in the Doppler axis due to the small crab angle.

Using a similar approach to [46], we inject one target into the range gate 290, which is located at broadside and Doppler bin 3 (corresponding to a Doppler frequency of 46.5 Hz) with a magnitude 0.003. The target location is

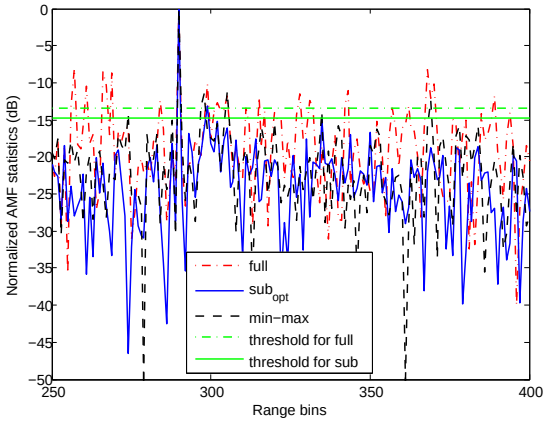

Fig. 7. The normalized AMF statistics versus range bins for MCRAM data at Doppler bin 3 and broadside.

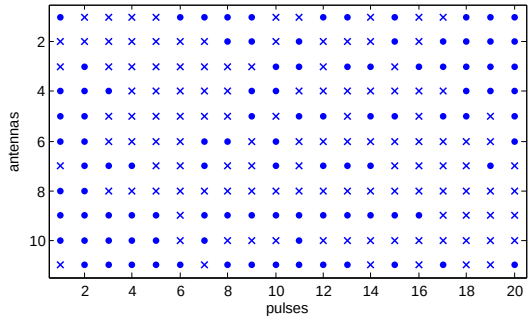

Fig. 8. The selected optimum sub-configuration with 100 antenna-pulse pairs.

also marked in Fig. 6. The selected optimum sub-configuration with 100 antenna-pulse pairs is shown in Fig. 8. The clutter spectrum corresponding to the optimum sub-configuration is presented in the right plot of Fig. 6. We can see that the clutter trajectory is stretched and narrowed, which increases the separation between the target and the clutter. The normalized AMF statistics versus range cells from 250 to 400 are shown in Fig. 7. The thresholds for both full and sub-configurations obtained from the simulations above are also shown. **Although one cannot ascertain whether all the peaks above the threshold are valid targets or false alarms, it is fair to say that it is highly unlikely that they would all be valid targets. It is certainly fair to say the optimum sub-configuration produces a much more reasonable detection performance, showing a significant reduction in the number of false alarms and indicating that there is only one target beside the one we injected. (Here needs editing)** The reason is that the maximum number of homogeneous training data is around 300 for acquisition 575 in flight 5 [45], whereas the full configuration requires 440 IID snapshots.

VI. Conclusion

This paper proposed a new perspective where we viewed the interleaved antenna-pulse data as forming a single large sensor array and investigated a thinned STAP by carrying out selection to this large sensor array. Each sensor, specified by the antenna number and pulse number, corresponds to an antenna-pulse pair. We reformulated Spatial and Spectral Correlation Coefficient (S^2C^2) to characterise the space-time separation between the target and clutter subspace. Based on the S^2C^2, we implemented the antenna-pulse selection as a maximization of the output SCNR and presented two algorithms to select the optimum antenna pulse pairs based on both Fourier basis and

eigenbasis. The performance of the proposed antenna-pulse selection strategy was validated using simulations and MCARM experimental data, which demonstrated its effectiveness on robust radar target detection while reducing the computational load and addressing the problems of clutter heterogeneity and limited sample support.

REFERENCES

[1] J. Ward, "Space-time adaptive processing for airborne radar," *Lincoln Laboratory Tech. Rept. ESC-TR-94-109*, Dec 1994.

[2] R. Klemm, *Principles of space-time adaptive processing*. The Institution of Electrical Engineers, 2002.

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

[4] J. Guerci, J. Goldstein, and I. Reed, "Optimal and adaptive reduced-rank STAP," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, no. 2, pp. 647–663, 2000.

[5] C. Richmond, "Performance of a class of adaptive detection algorithms in nonhomogeneous environments," *Signal Processing, IEEE Transactions on*, vol. 48, pp. 1248–1262, May 2000.

[6] I. Scott and B. Mulgrew, "Sparse LCMV beamformer design for suppression of ground clutter in airborne radar," *Signal Processing, IEEE Transactions on*, vol. 43, pp. 2843–2851, Dec 1995.

[7] M. Rangaswamy, J. H. Michels, and B. Himed, "Statistical analysis of the non-homogeneity detector for STAP applications," *Digital Signal Processing*, vol. 14, no. 3, pp. 253–267, 2004.

[8] E. Aboutanios and B. Mulgrew, "Hybrid detection approach for STAP in heterogeneous clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 3, pp. 1021–1033, 2010.

[9] D. J. Rabideau and A. Steinhardt, "Improved adaptive clutter cancellation through data-adaptive training," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 35, pp. 879–891, Jul 1999.

[10] J.-F. Degurse, L. Savy, and S. Marcos, "Reduced-rank STAP for target detection in heterogeneous environments," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 50, pp. 1153–1162, April 2014.

[11] E. Aboutanios and B. Mulgrew, "Heterogeneity detection for hybrid STAP algorithm," in *Radar Conference, 2008. RADAR '08. IEEE*, pp. 1–4, May 2008.

[12] W. Melvin and J. Guerci, "Knowledge-aided signal processing: a new paradigm for radar and other advanced sensors," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, pp. 983–996, July 2006.

[13] S. Bidon, O. Besson, and J.-Y. Tourneret, "Knowledge-aided STAP in heterogeneous clutter using a hierarchical Bayesian algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, pp. 1863–1879, July 2011.

[14] M. Steiner and K. Gerlach, "Fast converging adaptive processor or a structured covariance matrix," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 36, pp. 1115–1126, Oct 2000.

[15] A. Aubry, A. De Maio, L. Pallotta, and A. Farina, "Maximum likelihood estimation of a structured covariance matrix with a condition number constraint," *Signal Processing, IEEE Transactions on*, vol. 60, pp. 3004–3021, June 2012.

[16] T. Sarkar, H. Wang, S. Park, R. Adve, J. Koh, K. Kim, Y. Zhang, M. Wucjs, and R. Brown, "A Deterministic Least-Squares Approach to Space-Time Adaptive Processing (STAP)," *IEEE Transactions on Antennas and Propagation*, vol. 49, pp. 91–103, Jan. 2001.

[17] E. Aboutanios and B. Mulgrew, "A STAP algorithm for radar target detection in heterogeneous environments," in *Proc. IEEE/SP 13th Workshop on Statistical Signal Processing*, pp. 966–971, 2005.

[18] H. Deng, B. Himed, and M. Wicks, "Image feature-based space-time processing for ground moving target detection," *Signal Processing Letters, IEEE*, vol. 13, pp. 216–219, April 2006.

[19] A. Haimovich and Y. Bar-Ness, "An eigenanalysis interference canceler," *Signal Processing, IEEE Transactions on*, vol. 39, no. 1, pp. 76–84, 1991.

[20] C. H. Gierull, "Statistical analysis of the eigenvector projection method for adaptive spatial filtering of interference," *Radar, Sonar and Navigation, IEE Proceedings* -, vol. 144, pp. 57–63, Apr 1997.

[21] S. Sen, "OFDM radar space-time adaptive processing by exploiting spatio-temporal sparsity," *Signal Processing, IEEE Transactions on*, vol. 61, pp. 118 –130, jan.1, 2013.

[22] K. Sun, H. Meng, Y. Wang, and X. Wang, "Direct data domain STAP using sparse representation of clutter spectrum," *Signal Processing*, vol. 91, no. 9, pp. 2222–2236, 2011.

[23] Z. Yang, R. de Lamare, and X. Li, "Sparsity-aware space-time adaptive processing algorithms with l1-norm regularisation for airborne radar," *Signal Processing, IET*, vol. 6, no. 5, pp. 413–423, 2012.

[24] Z. Yang, R. De Lamare, and X. Li, "L1-regularized STAP algorithms with a generalized sidelobe canceler architecture for airborne radar," *Signal Processing, IEEE Transactions on*, vol. 60, pp. 674–686, Feb 2012.

[25] J. Ward, "Space-time adaptive processing with sparse antenna arrays," in *the Thirty-Second Asilomar Conference on Signals, Systems & Computers, 1998*, vol. 2, pp. 1537–1541 vol.2, 1998.

[26] E. Baranoski, "Sparse network array processing," in *Seventh IEEE Workshop on Statistical Signal and Array Processing*, pp. 145–148, 1994.

[27] X. Wang, E. Aboutanios, and M. Amin, "Reduced-rank STAP for slow-moving target detection by antenna-pulse selection," *Signal Processing Letters, IEEE*, vol. 22, pp. 1156–1160, August 2015.

[28] L. Brennan and F. Staudaher, "Subclutter visibility demonstration," *Adaptive Sensors, Tech. Rep., RL-TR-92-21*, 1992.

[29] F. Robey, D. Fuhrmann, E. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 28, pp. 208–216, Jan 1992.

[30] X. Wang, E. Aboutanios, and M. Amin, "Generalised array reconfiguration for adaptive beamforming by antenna selection," in *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on (accepted)*, May 2015.

[31] X. Wang, E. Aboutanios, M. Trinkle, and M. Amin, "Reconfigurable adaptive array beamforming by antenna selection," *IEEE Transactions on Signal Processing*, vol. 62, pp. 2385–2396, May 2014.

[32] J. Guerci and E. Baranoski, "Knowledge-aided adaptive radar at darpa: an overview," *Signal Processing Magazine, IEEE*, vol. 23, pp. 41–50, Jan 2006.

[33] L. Vandenberghe, S. Boyd, and S.-P. Wu, "Determinant maximization with linear matrix inequality constraints," *SIAM journal on matrix analysis and applications*, vol. 19, no. 2, pp. 499–533, 1998.

[34] R. Horst and N. V. Thoai, "DC programming: overview," *Journal of Optimization Theory and Applications*, vol. 103, no. 1, pp. 1–43, 1999.

[35] G. Gasso, A. Rakotomamonjy, and S. Canu, "Recovering sparse signals with a certain family of nonconvex penalties and DC Programming," *Signal Processing, IEEE Transactions on*, vol. 57, no. 12, pp. 4686–4698, 2009.

[36] A. Khabbazibasmenj and S. Vorobyov, "Robust adaptive beamforming for general-rank signal model with positive Semi-Definite constraint via POTDC," *Signal Processing, IEEE Transactions on*, vol. 61, no. 23, pp. 6103–6117, 2013.

[37] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[38] G. R. Lanckriet and B. K. Sriperumbudur, "On the convergence of the concave-convex procedure," in *Advances in neural information processing systems*, pp. 1759–1767, 2009.

[39] A. Beck, A. Ben-Tal, and L. Tetruashvili, "A sequential parametric convex approximation method with applications to nonconvex truss topology design problems," *Journal of Global Optimization*, vol. 47, no. 1, pp. 29–51, 2010.

[40] R. Horst, *Introduction to global optimization*. Springer, 2000.

[41] H. Tuy, *Convex analysis and global optimization*, vol. 22. Springer, 1998.

[42] Z.-Q. Luo, W.-K. Ma, A.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *Signal Processing Magazine, IEEE*, vol. 27, pp. 20–34, May 2010.

[43] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to algorithms*. MIT press, 2001.

[44] Y. Wu, J. Tang, and Y. Peng, "On the essence of knowledge-aided clutter covariance estimate and its convergence," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 1, pp. 569–585, 2011.

[45] B. Himed, "MCARM/STAP data analysis. volume II," tech. rep., DTIC Document, 1999.

[46] R. Adve and M. Wicks, "Joint domain localized processing using measured spatial steering vectors," in *Radar Conference, 1998. RADARCON 98. Proceedings of the 1998 IEEE*, pp. 165–170, May 1998.