ADAPTIVE RADAR DETECTION IN THE PRESENCE OF TEXTURED AND DISCRETE INTERFERENCE

A Dissertation
Presented to
The Academic Faculty

By

Jeong Hwan Bang

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in
Electrical and Computer Engineering

Georgia Institute of Technology

August 2013

ADAPTIVE RADAR DETECTION IN THE PRESENCE OF TEXTURED AND DISCRETE INTERFERENCE

Approved by:

Dr. Aaron D. Lanterman, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. William L. Melvin, Advisor
Georgia Tech Research Institute
Georgia Institute of Technology

Dr. James H. McClellan
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Douglas B. Williams
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Waymond R. Scott
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Brani Vidakovic
Wallace H. Coulter Department of Biomedical Engineering
Georgia Institute of Technology

Date Approved: June 23, 2013

ACKNOWLEDGEMENTS

This has been an incredible journey, where I have had the fortune of working with the brightest people in the industry. It was truly a humbling, yet rewarding experience. I am indebted to Dr. Melvin for the incredible amount of support he has provided me during my time at Georgia Tech. He has gone above and beyond to help me through the many challenges I've faced in pursuing my doctorate degree. Dr. Melvin's knowledge in the subject of radar signal processing is inspiring and I am very fortunate to have learned from one of the leaders in the field. I thank Dr. Lanterman for his continual support and openness to the research I wished to pursue. He has provided me with connections that helped me through my thesis and connections that will pave the way for my future career goals.

I thank my family and friends at home for their understanding and support. The friends I have met at Georgia Tech, especially, made it a memorable experience. Looking back, I enjoyed the times we studied together, even towards the daunting preliminary exam. It was great to have peers challenge you to think critically and differently.

Last and mostly, I would like to thank my mother for her never-ending support. She challenged me to aspire for higher achievements in life and I would have never reached this far without her blessing. Even as I stumbled on the nadir of my life, I knew I'd persevere, as I could never let her down.

TABLE OF CONTENTS

LIST OF TABLES

PAGE

LIST OF FIGURES

PAGE

SUMMARY

Under a number of practical operating scenarios, traditional moving target indicator (MTI) systems inadequately suppress ground clutter in airborne radar systems. Due to the moving platform, the clutter gains a nonzero relative velocity and spreads the power across Doppler frequencies. This obfuscates slow-moving targets of interest near the "direct current" component of the spectrum. In response, space-time adaptive processing (STAP) techniques have been developed that simultaneously operate in the space and time dimensions for effective clutter cancellation. STAP algorithms commonly operate under the assumption of homogeneous clutter, where the returns are described by complex, white Gaussian distributions. Empirical evidence shows that this assumption is invalid for many radar systems of interest, including high-resolution radar and radars operating at low grazing angles. When the heterogeneity stems from textured clutter, we demonstrate that a 0.5 loss in detection probability can be expected for a system that maintains a false alarm rate of 1×10^{-4}. Similarly, a 0.3 loss in detection probability can be expected in the presence of point clutter. We are interested in heterogeneous cases, i.e., cases when the Gaussian model no longer suffices [1],[2].

Hence, the development of reliable STAP algorithms for real systems depends on the accuracy of the heterogeneous clutter models. The clutter of interest in this work includes heterogeneous texture clutter and point clutter. We have developed a cell-based clutter model (CCM) that provides simple, yet faithful means to simulate clutter scenarios for algorithm testing [2]. The scene generated by the CMM can be tuned with two parameters, essentially describing the spikiness of the clutter scene. In one extreme, the texture resembles

point clutter, generating strong returns from localized range-azimuth bins. On the other hand, our model can also simulate a flat, homogeneous environment. We prove the importance of model-based STAP techniques, namely knowledge-aided parametric covariance estimation (KAPE), in filtering a gamut of heterogeneous texture scenes. We demonstrate that the efficacy of KAPE does not diminish in the presence of typical spiky clutter. In addition, for a system that maintains a false alarm rate of 1×10^{-4}, we show that KAPE actually improves the detection probability by 0.1 compared to standard algorithms in homogeneous clutter.

Computational complexities and susceptibility to modeling errors prohibit the use of KAPE in real systems. The computational complexity is a major concern, as the standard KAPE algorithm requires the inversion of an $MN \times MN$ matrix for each range bin, where M and N are the number of array elements and the number of pulses of the radar system, respectively. This is compounded with the requirement of greater than MN power estimates, one for each steering vector. We developed a Gram Schmidt (GS) KAPE method that circumvents the need of a direct inversion and reduces the number of required power estimates. We demonstrate that no losses in detection, nor SINR, are incurred as a result of the new methods [3]. These developments allow for real-system implementations of the KAPE algorithm.

Another unavoidable concern is the performance degradation arising from uncalibrated array errors. This problem is exacerbated in KAPE, as it is a model-based technique; mismatched element amplitudes and phase errors amount to a modeling mismatch. We demonstrate that SINR losses of -35 dB can be expected if array errors are not accounted for. Current calibration algorithms can boost performance, but can still incur losses of -10 dB. We have

developed the power-ridge aligning (PRA) calibration technique, a novel iterative gradient descent algorithm that can maintain a -1 dB SINR loss in the presence of array errors, a vast improvement over the current methods [3].

Chapter 1

INTRODUCTION

The aim of moving target indicator (MTI) radar is to extract moving targets from stationary clutter by processing Doppler shifts from a sequence of temporal pulse returns [4]. By relating the phases from successive pulses, the differential phases can be measured to calculate the frequency shift. For a moving target, the received frequency f_r is related to the transmit frequency f_t by

$$f_r = \left(\frac{1 + 2v_r/c}{1 - 2v_r/c}\right) f_t \approx (1 + 2\, v_r/c) f_t \,,$$

where v_r and c are the radial velocity and the speed of light, respectively. The approximation reasonably assumes that $v_r \ll c$. In contrast, for a stationary radar system, ground clutter has a zero relative radial velocity and does not accrue phase over time. This difference allows us to separate clutter from moving targets via the Doppler spectrum, as ground clutter tends to center around the "direct current" (DC) component. Targets with faster radial velocities attain greater separability, accruing larger phase differences over time.

We are interested in airborne MTI systems, where the radar platform has a nonzero velocity. In this case, ground clutter gains a velocity proportional to the velocity of the aircraft. This adds complexity to the problem, as the shifting and spreading of the clutter spectrum can potentially obfuscate the targets of interest.

This led to the development of space-time adaptive processing (STAP) algorithms designed to mitigate clutter on airborne radar systems [4], [5]. Building on the MTI concept of using time-domain samples, STAP also uses spatial samples collected via an array of M antenna elements, allowing for two-dimensional space-time filters. Additionally, the array gives spatial directivity to the receiver through beamforming. Figure 1 illustrates an example space-time spectrum of a clutter-plus-noise signal on a uniform linear array (ULA) system with $M = N = 16$, where $0°$ spatial axis corresponds to the beam steered perpendicular to the platform direction. The clutter aligns on a clutter ridge, the diagonal of the 2D spectrum. When the array is steered at $0°$, the received clutter return has zero velocity because it is moving in parallel with the platform. The clutter starts to accrue radial velocity as the beam is steered away from $0°$, hence explaining the clutter ridge structure and the innate coupling of the Doppler shifts and angle of arrivals (AOA). A single element, stationary MTI system corresponds to the $0°$ cut. As discussed, the clutter is centered around the DC component of the Doppler spectrum.

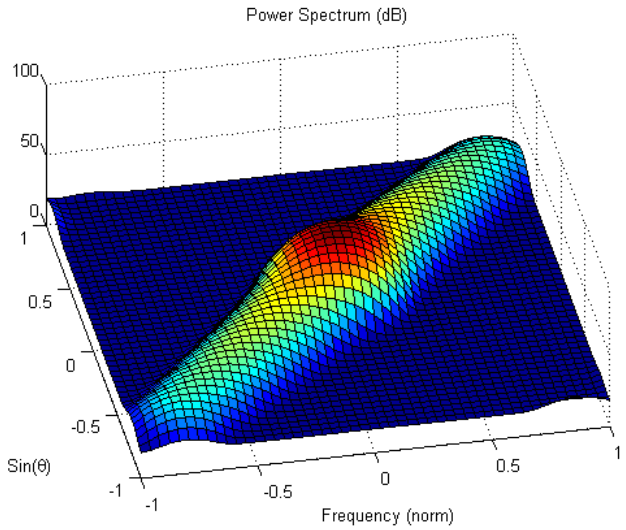


Figure 1. Example angle-Doppler spectrum for airborne radar clutter.

STAP is designed to exploit the space and time dependencies of ground clutter to extract targets of interest. For a given range R_o, the radar system receives N return pulses on an array of M elements. Upon designating an arbitrary element as the reference, the other elements of the array receive a positive or negative delayed version of the source signal due to their spatial differences. Time delays correspond to phase shifts in the frequency domain. Hence, we define the spatial steering vector $\boldsymbol{s}_s(\phi, \theta)$ as the collection of phases across the array at an assumed steered direction defined by azimuth angle ϕ and elevation angle θ. The first element is typically used as the reference assigned a phase of zero. Assuming a ULA with element spacing d and design wavelength λ_o,

$$\boldsymbol{s}_s(\phi, \theta) = \left[1 \quad \exp\left\{j\frac{2\pi d}{\lambda_o}\cos(\theta)\sin(\phi)\right\} \quad \cdots \quad \exp\left\{j(M-1)\frac{2\pi d}{\lambda_o}\cos(\theta)\sin(\phi)\right\}\right]^T.$$

We can similarly define the temporal steering vector $\boldsymbol{s}_t(v_r)$ as the collection of progressive phases along the received pulses for a target with radial velocity v_r. Denoting f_{PRF} as the pulse repetition frequency (PRF), i.e., the rate at which the pulses are pinged,

$$\boldsymbol{s}_t(v_r) = \left[1 \quad \exp\left\{j\frac{4\pi v_r}{\lambda_o f_{PRF}}\right\} \quad \cdots \quad \exp\left\{j(N-1)\frac{4\pi v_r}{\lambda_o f_{PRF}}\right\}\right]^T.$$

The $MN \times 1$ space-time steering vector $\boldsymbol{s}(v_r, \phi, \theta)$ is formed by combining the spatial and temporal steering vectors

$$\boldsymbol{s}(v_r, \phi, \theta) = \boldsymbol{s}_t(v_r) \otimes \boldsymbol{s}_s(\phi, \theta),$$

where $\otimes$ is the Kronecker product. The received data can be compactly visualized as a matrix $\boldsymbol{X}$, where an element $[\boldsymbol{X}]_{m,n}$ corresponds to the data of the m^{th} array element and n^{th} pulse. For STAP calculations, it is convenient to stack the matrix into a single vector

$$\boldsymbol{x} = \mathrm{col}(\boldsymbol{X}),$$

where $\mathrm{col}(\cdot)$ corresponds to the column operation that reshapes a matrix into a vector.

1.1. Optimal Filter

The aim of STAP is to suppress unwanted clutter from the received $MN \times 1$ data vectors $\boldsymbol{x}_l,\ l = 1, \ldots L$ with a finite impulse response (FIR) filter specified by $\boldsymbol{w}$, where L is the number of received range bins. The received data can be written as

$$\boldsymbol{x} = \boldsymbol{x}_s + \boldsymbol{n},$$

which is the sum of the desired signal $\boldsymbol{x}_s$ and clutter-plus-noise $\boldsymbol{n}$. Hence, the STAP output is obtained via the dot product with the weights vector:

$$\boldsymbol{z} = \boldsymbol{w}^H \boldsymbol{x}.$$

To obtain the optimal filter, we can calculate the weight vector $\boldsymbol{w}$ that will yield the maximum signal-to-noise ratio (SNR). The resulting noise filter, and hence the optimum STAP filter, for data received at steering angle (ϕ, θ) with radial velocity v_r is the Wiener-Hopf filter [6]

$$\boldsymbol{w} = \alpha \boldsymbol{R}^{-1} \boldsymbol{s}(v_r, \phi, \theta),$$

where $\boldsymbol{R}$ denotes the exact clutter-plus-noise covariance matrix and α is a complex scaling term. We refer to the exact covariance matrix $\boldsymbol{R}$ as the clairvoyant covariance matrix, as it is not known in practical applications and must be estimated. Assuming we have $L \geq MN$ samples of the data, a simple approach is to compute the empirical correlation matrix of the data:

$$\widehat{\boldsymbol{R}} = \frac{1}{L} \sum_{l=1}^{L} \boldsymbol{x}_l \boldsymbol{x}_l^H,$$

often referred to as the sample matrix inversion (SMI) method. Given that the clutter is a white Gaussian process and that the set of samples $\boldsymbol{x_1}, \ldots, \boldsymbol{x}_L$ share the same statistics, SMI suffers, on average, a 3 dB loss from the clairvoyant case when using $L = 2MN - 3$ range samples [7]. This scenario effectively describes performance in homogeneous clutter. Unfortunately, these assumptions are overly optimistic in many situations, leading to poor performance. For example, it has been observed that high-resolution and low-grazing angle radars exhibit clutter that is spikier [8],[9]. As we will show, when the heterogeneity stems from textured clutter with a single-channel, single-pulse clutter-to-noise-ratio of 25 dB, we can expect a 0.5 loss in detection probability for a system that maintains a false alarm rate of 1×10^{-4}. Furthermore, a 0.3 loss in detection probability has been observed with the inclusion of point clutter in a similar setting. The nature of heterogeneous texture clutter and point clutter will be further elaborated in Chapter 2. These losses necessitate the specialized algorithms we have developed.

1.2. Cell-Based Clutter Model

The losses we demonstrate underline the importance of properly modeling the clutter scenes of interest. The performance outcomes of our developed algorithms are only meaningful if the simulated clutter matches the real clutter scenes we aim to filter. We have developed the cell-based clutter model (CCM), a phenomenological approach to clutter modeling. Specifically, it is based on a compound model of clutter where two underlying clutter distributions are assumed: the modulation of the fast moving and the slow moving components of the scene. The CCM demonstrates its versatility by being able to simulate homogeneous clutter, heterogeneous texture clutter, and point clutter for algorithm analysis. This is elaborated in Chapter 3. The CCM is designed to be radar independent, only requiring parameters of the scene itself. This allows flexibility in comparing performances of different radar systems, such as systems with varying operational frequencies and array configurations.

Chapter 3 illustrates the advantage of applying model-based STAP techniques in the presence of heterogeneous texture clutter. Specifically, we demonstrate the effectiveness of the knowledge-aided parametric covariance estimation (KAPE) in handling a range of spiky, heterogeneous clutter scenes. Under our simulation parameters, KAPE improves the detection probability by 0.6 over the ubiquitous SMI algorithm for a system that maintains a false alarm rate of 1×10^{-4}. This is attributed to the KAPE's ability to adapt over a single range bin, whereas the SMI averages data over the range swath. In general, this is the fundamental reason for the poor adaptivity of conventional methods in spiky clutter.

1.3. Knowledge-aided Parametric Covariance Estimation

Knowledge-aided parametric covariance estimation (KAPE) is a model-based approach to STAP that parameterizes the clutter ridge of the clutter covariance matrix. Its ability to adapt over single range bins, namely the cell-under test (CUT), makes it an ideal candidate to combat spiky heterogeneous clutter. We describe KAPE in more detail in Section 3.2.

Despite KAPE's proficiency in adapting to clutter scenes of interest, it is burdened by large computational requirements that prohibit its use in real systems. The crux of the computations are the $MN \times MN$ matrix inversions required to calculate the STAP weights. Since KAPE operates on single range bins, the inversion must be performed for every range bin within the range swath. As discussed in Section 3.2, KAPE also requires power estimates for each of the steering vectors that form the clutter ridge. To address this problem, we developed a Gram Schmidt (GS) KAPE algorithm that bypasses the need to calculate direct inversions, making it feasible for real-time system implementation. Furthermore, GS KAPE reduces the number of steering vectors in its covariance matrix estimate, hence lessening the number of required power estimates. Section 4.2 details this novel algorithm.

In contrast to SMI techniques, KAPE is more susceptible to array errors, since it is a model-based technique. Uncalibrated array errors amount to a model mismatch; SINR losses of -35 dB can be expected, as demonstrated in Section 4.3. Some array calibration algorithms can ameliorate such losses, but they can still incur SINR losses up to -10 dB. These heavy losses may restrict KAPE's effectiveness in real systems. To address this concern, we have developed the power-ridge aligning (PRA) calibration technique, a novel iterative gradient descent approach. Using this PRA technique, we can maintain a -1 dB SINR loss in the

presence of array errors, a significant boost over existing methods. Section 4.3 describes the approach and compares its performance.

We refer to enhanced-KAPE (E-KAPE) as the collective use of the various upgrades we have developed over the standard KAPE algorithms. To demonstrate the performance improvements, Section 4.4 assesses it over a various scenarios, including the presence of point clutter and array errors. Coupled with the computational reduction, E-KAPE maintains the feasibility of real system implementation.

Chapter 2

BACKGROUND

This chapter aims to elucidate the contributions made in this dissertation by presenting the relevant background information. Section 2.1 begins by defining the assumed radar signal model for our research. We then clarify the distinctions between homogeneous and heterogeneous clutter models in Section 2.2 and Section 2.3. Models for targets of interest are presented in Section 2.4, which include the Swerling 1 model that we incorporate in our simulations. Section 2.5 focuses on important detection statistics that are directly used in our algorithm performance studies. This is followed by a survey of methods to reduce their computational complexities. Section 2.6 elaborates on the knowledge-aided parametric covariance estimation (KAPE) method, which is central to the contributions in this dissertation. Finally, Section 2.7 discusses the performance metrics that are employed to compare STAP algorithms.

2.1. Radar Returns

The radar return over an area of interest is often approximated as the combination of the returns from many individual scatterers. The transmitted narrowband radar pulse $x_T(t)$ can be generally written as

$$x_T(t) = a(t)\sin[2\pi f_T t + \theta(t)],$$

where f_T is the transmit frequency, $a(t)$ is the amplitude modulation term and $\theta(t)$ is the phase modulation term. Assuming the scene consists of a single point source at range $R_o = ct_o/2$, where c is the speed of light, the received pulse $x_R(t)$ is a delayed and attenuated echo of the transmitted pulse [4]:

$$x_R(t) = \tilde{a}(t - t_o)\sin\left[2\pi f_T(t - t_o) + \tilde{\theta}(t - t_o)\right] + n(t),$$

where $\tilde{a}(t)$ is the attenuated amplitude due to propagation loss, $\tilde{\theta}(t)$ is the distorted phase modulation, and $n(t)$ is the receiver noise. Phase distortion can be attributed to factors such as Doppler shifts. In STAP processing, the signal is received over an array of M spatial elements and observed over N pulses. Hence, we receive MN signals of varying delays that are used to improve the SINR of the system and provide directivity, as explained in Chapter 1.

Let P_T be the power of the transmitted signal, g be the array gain, and A_e be the antenna aperture area. Then received power is

$$P_R = \frac{P_T g A_e \sigma}{(4\pi)^2 R_o^4} \text{ Watts},$$

where σ is the radar cross section (RCS) that describes the fraction of the reflected transmit power. Hence, the RCS values are specific to the range R_o. In our case, we are interested in scatterers in an illuminated area A, where we assume that the individual scatterers share the same statistics. The reflected power is then quantified by the RCS in unit area, which has units of m^2 [4].

Due to the intractable amount of scatterers, the RCS value is highly sensitive to the aspect angle and the transmit frequency of the radar. Hence, in practice, a statistical description of the RCS value is adopted to describe the scene. Although this model removes the azimuth-aspect dependencies from scatterer fluctuations, the grazing angle can still affect the outcome RCS [10]. At low grazing angles, shadowing and multipath become a problem, while returns from high grazing angles are plagued with facet reflections. To account for these interferences, we adopted a constant gamma model [11],[12]. This model describes the reflectivity as

$$\sigma^o(\psi) = \gamma \sin\psi,$$

where γ is the normalized reflectivity and ψ is the grazing angle. The RCS is then calculated as $\sigma_c = A_c\sigma^o$, where A_c is the area of radar illumination.

To model the clutter signal received by the radar system, we divide the iso-range ring of interest into N_p clutter patches, each patch corresponding to an azimuth interval. The range interval of the ring depends on the sampling rate of the receiver. For this exposition, we assume a fixed elevation angle. Let $\boldsymbol{s}_k$ denote the space-time steering vector pointed towards the k^{th} clutter patch. Then the received data $\boldsymbol{x}$ is

$$\boldsymbol{x} = \sum_{k=1}^{N_p} \alpha_k \boldsymbol{s}_k,$$

where α_k is the complex gain from the k^{th} clutter patch, which also accounts for the array gain as well as the reflectivity. This approximation forms the basis for the STAP algorithms

of interest in this dissertation. We see that the steering vectors $\boldsymbol{s}_k$ correspond exactly to points along the clutter ridge, the region STAP is designed to suppress.

2.2. Homogeneous model

If we assume that the scatterers contributing to the reflectivity γ in the area of interest consist of independent and identically distributed (i.i.d.) components, we can invoke the Central Limit Theorem (CLT), which states that the real and imaginary components of the composite voltage will be well approximated by white Gaussian random variables [13]. That is, the reflectivity values are drawn from the complex Gaussian distribution:

$$\gamma \sim CN[\mu, \sigma^2],$$

with mean μ and variance σ^2. It follows that the magnitudes $|\gamma|$ are Rayleigh distributed. These assumptions define the homogeneous clutter model. Under such conditions, applying the sample matrix inverse (SMI) for clutter cancellation results in a 3 dB SINR loss, on average, from the clairvoyant case when using $L = 2MN - 3$ range samples [7]. Hence SMI provides us with an effective and efficient clutter filter when the homogeneous criteria are met.

2.3. Heterogeneous Clutter

The homogeneous assumption may be too ideal to adequately describe scenes of interest. For example, spurious spikes observed from low grazing angles violate the i.i.d. assumption of the scatterers [14]. In addition, received data from high-resolution radar is poorly modeled with the homogeneous model, where the area of illumination is smaller; a small number of scatterers do not satisfy the assumptions of the CLT as well as a large number of scatterers. In general, heavier distribution tails have been observed that are inadequately parameterized by Gaussian models [15],[16]. Numerous distributions have been tried to model the deviations, including Weibull and log-normal distributions [16],[17]. For this dissertation, we are particularly interested in the K-distribution, noted for its adeptness in modeling various states of sea clutter [18].

2.3.1. Texture Clutter

The purpose of heterogeneous clutter models is to better parameterize clutter distributions that do not satisfy the homogeneous conditions. The compound clutter model is a heterogeneous clutter model that attempts to match the underlying phenomenology by modulating two distributions [19],[20]. In particular, the two processes of interest characterize the scatterers with long time-correlations, or the slow-moving components, and the scatterers with short time-correlations. We can view the homogeneous model as the special case where only the short time-correlated scatterers exist, justifying the independence of the scatterer statistics. For sea clutter, the sea state dictates the long time-correlated clutter, modulated by the impulsive capillary waves [9]. Of the various cases of compound

clutter models, we are interested in the K-distribution model, as it has shown to be a good empirical fit to real data and presents a relatively simple parameterization [8],[18].

The K-distributed model uses a gamma random variable to model the long time-correlated process, modulated with a white complex Gaussian process describing the fast moving components. Specifically, we can write the probability density function for a K-distributed random variable as

$$f(x) = \int_0^\infty f(x|\gamma) f(\gamma) d\gamma\,,\ x \geq 0,$$

where $f(x|\gamma)$ is the Rayleigh probability density function

$$f(x|\gamma) = \frac{x}{\gamma} \exp\left(\frac{-x^2}{2\gamma}\right), \qquad x \geq 0$$

and $f(\gamma)$ is the gamma density function

$$\Gamma(\gamma|\alpha,\beta) \equiv f(\gamma) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \gamma^{\alpha-1} \exp\left(\frac{-\gamma}{\beta}\right), \qquad \gamma > 0,$$

with shape and scale parameters α and β, respectively, and the gamma function $\Gamma(\cdot)$. Although the K-distribution is only parameterized with two parameters, it has the flexibility to model the clutter heterogeneities of interest. The gamma ratio $r_\gamma = E[\gamma]/\mathrm{std}[\gamma]$ aptly describes the roughness of the scene [21]. For example, a low value of r_γ trends toward a spikier scene, as it implies an increase in the standard deviation. In the other case, noting that $E[\gamma] = \alpha\beta$, we see that the shape parameter increases with increasing r_γ. The Gamma distribution converges to a Gaussian distribution as $\alpha \to \infty$ [13], and hence yielding a milder

scene. Figure 2 plots the empirical K-distributions for various gamma ratios with the mean fixed at $E[\gamma] = 1$. We observe that the distributions trend towards spikier scenes with decreasing r_γ.

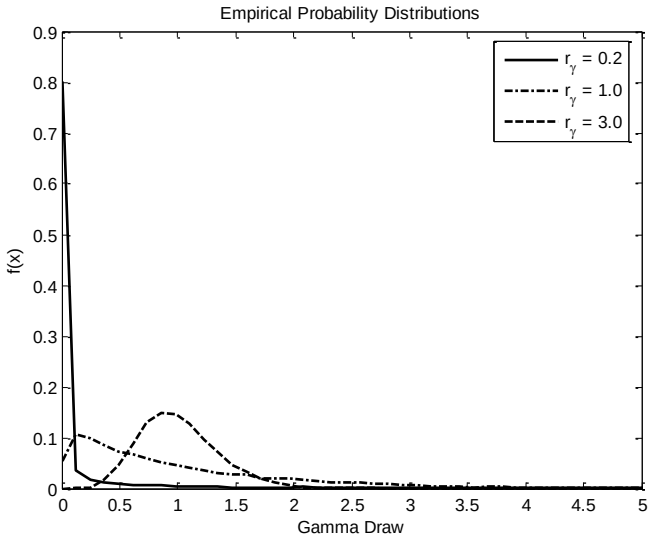


Figure 2. Gamma distributions for various gamma ratio values.

2.3.2. Point Clutter

The K-distribution describes the texture of the scene, as it models the contribution of the many scatterers in the region. It does not, however, model the highly reflective, outlier objects in the area of interest. These include speeding vessels and stationary objects such as docked boats. The received data from these discretes are localized and generally impulsive, spanning a fraction of the azimuth and range resolution of the system. We refer to these heterogeneities as point clutter.

There are two distinct classes of point clutter: stationary clutter discretes (CD) and targets in secondary data (TSD) [22]. Each of these pose a different problem to STAP. CDs are

inherently stationary, and hence reside along the clutter ridge of the data. If a discrete is not properly nulled, the residue clutter power may spread across the Doppler bins, triggering false alarms. On the other hand, TSDs are seen away from the clutter ridge due to their nonzero radial velocities. When TSDs are present in the training data, e.g., the range interval to average over as in the case of SMI, nulls may be unintentionally placed in their location in the Doppler-angle spectrum. These nulls may reduce the detection probability, since they will filter out the targets of interest.

To maintain the impulsiveness, data discretes are simulated by seeding into the received data rather than accounting for them in the RCS estimation. Let $\boldsymbol{x}_k$ be the data received from the k^{th} range bin. Let $\boldsymbol{s}_{CD}$ be the steering vector corresponding to a CD. Then the injected data $\boldsymbol{x}'_k$ is formulated as

$$\boldsymbol{x}'_k = \boldsymbol{x}_k + \alpha_{CD}\boldsymbol{s}_{CD},$$

where α_{CD} is the corresponding gain. TSDs are modeled in a similar manner:

$$\boldsymbol{x}'_k = \boldsymbol{x}_k + \alpha_{TSD}\boldsymbol{s}_{TSD},$$

where α_{TSD} and $\boldsymbol{s}_{TSD}$ are the TSD gain and steering vector, respectively.

2.4. Targets

Target data is simulated by injection, similar to the insertion of point clutter. In this work, we determine the power of the injected target by adopting the well-established Swerling models

[4],[23]. The Swerling models were chosen as a basis for our performance analysis because they form the foundation of much experimental and theoretical work in the radar literature. They were initially proposed to statistically describe the returns over multiple pulses within the cell-under-test (CUT). In particular, the Swerling 1 and Swerling 2 models describe scenarios where the return target signal is composed of many independent and identically distributed scatterers, which is in alignment with our initial assumptions of the scene. The magnitude of the returns are described with the Rayleigh distribution:

$$p(x) = \frac{1}{\sigma}\exp(-x/\sigma),$$

which has a variance of σ^2. The Swerling 1 model further assumes that the returns vary stochastically between the coherent processing intervals (CPI), but remain constant between the pulses of the CPIs. This is appropriate for an object with long temporal correlation times. On the other hand, the Swerling 2 model assumes that the returns vary stochastically between all pulse returns, so it is appropriate for an object with short temporal correlation times. Other Swerling models examine the case when the magnitudes of the returns are dominated by a strong scatterer, resulting in a Chi-squared distribution:

$$p(x) = \frac{4x}{\sigma^2}\exp\left(-\frac{2x}{\sigma}\right),$$

which has a variance of $\sigma^2/2$. In total, there are four distinct Swerling models as listed in Table 1. For our simulations, we will strictly assume a Swerling 1 target.

Table 1. Description of the four common Swerling models.

	Short Correlation Times	**Long Correlation Times**
Uniform Scatterers	Swerling 1	Swerling 2
Dominant scatterer	Swerling 3	Swerling 4

2.5. Detectors

For targets with Doppler and angle corresponding to the steering vector $\boldsymbol{s}$, the optimal filter is given by the Wiener-Hopf equation:

$$\boldsymbol{w} = \boldsymbol{R}^{-1}\boldsymbol{s},$$

where $\boldsymbol{R}$ is the clairvoyant covariance matrix. In practice, neither the exact steering vector nor the covariance matrix are known. Array errors deviate the spatial steering vector from the uniform linear array model discussed in Section Chapter 1. We denote the corrected steering vector as $\boldsymbol{v}$ and $\widehat{\boldsymbol{R}}$ as the covariance matrix estimate. The corrected steering vector $\boldsymbol{v}$ accounts for the channel amplitude and phase errors caused by practical implementation issues, such as element displacements and mutual coupling. If we denote the $M \times 1$ vector of exact array errors as $\boldsymbol{\epsilon}$, then the corrected steering vector is

$$\boldsymbol{v} = \boldsymbol{s} \odot (\mathbf{1}_N \otimes \boldsymbol{\epsilon}),$$

where $\mathbf{1}_N$ is an $N \times 1$ vector of ones. As noted in Section 1.1, the sample matrix inverse (SMI) gives a good estimate in homogeneous clutter and will be the main benchmark of our

algorithms. Recall that the covariance matrix estimate is essentially an average of the outer products of the given data:

$$\widehat{\boldsymbol{R}} = \frac{1}{L}\sum_{l=1}^{L} \boldsymbol{x}_l \boldsymbol{x}_l^H,$$

where L is the number of available range samples. The following subsections will focus on direct and indirect modifications to this estimate.

Ultimately, we require a statistic derived from the filter output that we can compare to a threshold. The threshold, which may be fixed or adaptive, will determine whether a target exists in the CUT. This implies a normalization of the outputs of the filter $\boldsymbol{w}$ to account for varying background noise. Some commonly used algorithms include Kelly's generalized likelihood ratio (GLR) detector [24] and the adaptive coherence estimator (ACE) [25]. Kelly's GLR approaches the problem as a hypothesis testing problem with two hypotheses: the noise-only case and the signal-plus-noise case. This approach is motivated by the difficulty encountered in choosing appropriate thresholds to maintain a desired false alarm rate using the SMI filter. Assuming unknown noise statistics and signal amplitude, the GLR is calculated in which the unknown parameters are substituted with their maximum-likelihood estimates. The detection statistic is

$$z = \frac{\left|\boldsymbol{v}^H \widehat{\boldsymbol{R}}^{-1} \boldsymbol{x}\right|^2}{\boldsymbol{v}^H \widehat{\boldsymbol{R}}^{-1} \boldsymbol{v} \left[1 + \frac{1}{L} \boldsymbol{x}^H \widehat{\boldsymbol{R}}^{-1} \boldsymbol{x}\right]}.$$

Kelly's GLR is compared to a threshold $L\eta$, where L is the number of available range samples and η is chosen to maintain a desired false alarm rate. Note that the threshold directly depends on the support size.

Another useful detector is the adaptive coherence estimator (ACE) given by

$$z = \frac{\left|\boldsymbol{v}^H \widehat{\boldsymbol{R}}^{-1} \boldsymbol{x}\right|^2}{\boldsymbol{v}^H \widehat{\boldsymbol{R}}^{-1} \boldsymbol{v}^H \left(\boldsymbol{x}^H \widehat{\boldsymbol{R}}^{-1} \boldsymbol{x}\right)} .$$

ACE was designed to be invariant with respect to scaling inconsistencies between the CUT and the rest of the range samples. It is essentially the angle between the CUT and the steering vector projected onto the subspace defined by $\widehat{\boldsymbol{R}}^{-1}$, or

$$\frac{\left|\langle \widehat{\boldsymbol{R}}^{-1/2} \boldsymbol{v}, \widehat{\boldsymbol{R}}^{-1/2} \boldsymbol{x} \rangle_2^2\right|^2}{\left\|\widehat{\boldsymbol{R}}^{-1/2} \boldsymbol{v}\right\|_2^2 \left\|\widehat{\boldsymbol{R}}^{-1/2} \boldsymbol{x}\right\|_2^2} .$$

Hence, ACE can better filter out false detections away from the steered direction.

The adaptive matched filter (AMF) is an efficient detection statistic that also approaches the problem as a hypothesis test [26]. The AMF follows from Kelly's GLR for $L \to \infty$:

$$z = \frac{\left|\boldsymbol{v}^H \widehat{\boldsymbol{R}}^{-1} \boldsymbol{x}\right|^2}{\boldsymbol{v}^H \widehat{\boldsymbol{R}}^{-1} \boldsymbol{v}} .$$

We immediately see the reduction in computational complexity compared to Kelly's GLR and ACE by the omission of the matrix operation $\boldsymbol{x}^H \widehat{\boldsymbol{R}}^{-1} \boldsymbol{x}^H$. AMF is widely used in practice for its simplicity and effectiveness. Likewise, the AMF is used for our simulations to calculate performance metrics for algorithm comparisons.

2.5.1. Dimension Reduction

One common requirement of the detectors discussed above is an estimate of the inverse covariance matrix $\widehat{\boldsymbol{R}}^{-1}$. When the SMI is employed, we require at least MN range samples to ensure a full-rank covariance matrix $\widehat{\boldsymbol{R}}$ and $2MN$ to maintain 3 dB SINR loss on average [7]. Even for our modest test system with $M = 6$ channels and $N = 32$ pulses, this implies a minimum of $L = 192$ range samples. The SMI further requires that the samples used for the estimation maintain homogeneity over the training interval, which is an optimistic assumption.

Several dimension reducing techniques have been proposed to address this problem. These include a variety of eigenvalue techniques that assume the rank of the clutter component is much lower than MN [27],[28],[29]. They maintain that the clutter has a rank r_c and attribute the full rank of the covariance matrix to the white background noise. The covariance estimate is then reformulated with a subset of the r_c eigenvectors and eigenvalues, assuming that the principal components are directed towards the clutter. This is a valid assumption, as the clutter components typically overpower the background noise. These algorithms face the additional problem of estimating the appropriate value of r_c. Since the eigenvectors are orthogonal, the inverse covariance matrix can be derived without a direct matrix inversion. Specifically, if we denote $\boldsymbol{Q}$ as the matrix of the r_c eigenvectors and $\boldsymbol{\Lambda}$ as the $r_c \times r_c$ diagonal matrix with corresponding eigenvalues, the inverse of the clutter component is calculated as

$$\widehat{\boldsymbol{R}}^{-1} = \boldsymbol{Q}\boldsymbol{\Lambda}^{-1}\boldsymbol{Q}^{H},$$

where $\boldsymbol{\Lambda}^{-1}$ is trivially calculated by taking the inverse of each diagonal entry of $\boldsymbol{\Lambda}$. Hence, these approaches do not require MN range samples.

The extended factor algorithm (EFA) simplifies the problem by only processing $K < N$ adjacent Doppler bins at a time instead of simultaneously processing all data within the CPI. This essentially branches STAP into N parallel filter banks, each with dimension KM [30]. The number of adjacent bins K is chosen to be big enough to capture most of the correlated information in the Doppler domain and small enough to sufficiently reduce the STAP dimension. The Joint-Domain Localized (JDL) method takes this idea a step further by localizing in the angle space as well as the Doppler space [31]. This method can exploit the fact that clutter lies along the clutter ridge in the angle-Doppler spectrum. Both EFA and JDL significantly reduce computational burden, since covariance matrix inversion is $O(\mathrm{DOF}^3)$, where DOF is the degrees of freedom.

We are particularly interested in the EFA algorithm for its effectiveness and simplicity. Let $\boldsymbol{F} = [\boldsymbol{f}_0, \boldsymbol{f}_1, \dots, \boldsymbol{f}_{N-1}]$ be the $N \times N$ DFT matrix. We initially apply a Hamming window $\boldsymbol{w}_{hamm}$ of length N for sidelobe suppression, resulting in

$$\boldsymbol{F}' = [\boldsymbol{f}'_0, \boldsymbol{f}'_1, \dots, \boldsymbol{f}'_{N-1}] = [\boldsymbol{w}_{hamm} \odot \boldsymbol{f}_0, \boldsymbol{w}_{hamm} \odot \boldsymbol{f}_1, \dots, \boldsymbol{w}_{hamm} \odot \boldsymbol{f}_{N-1}],$$

which can compactly computed as

$$\boldsymbol{F}' = \boldsymbol{w}_{hamm}^T \mathbf{1}_N \odot \boldsymbol{F}.$$

Assuming K is odd, the n^{th} filter bank of EFA can be represented with a transformation matrix

$$\boldsymbol{T}_n = \left[\boldsymbol{f}'_{n-(K-1)/2}, \cdots, \boldsymbol{f}'_{n-1}, \boldsymbol{f}'_{n}, \boldsymbol{f}'_{n+1}, \cdots, \boldsymbol{f}'_{n+(K-1)/2}\right] \otimes \boldsymbol{I}_M \,,$$

where $\boldsymbol{I}_M$ is an $M \times M$ identity matrix. The cyclic property of the DFT is used to represent $\boldsymbol{f}_i$ when $i < 0$ or $i > N-1$. When K is even, the filters are frequency shifted to maintain the symmetry of the transformation:

$$\boldsymbol{T}_n = \left[\boldsymbol{f}'_{n-(K+1)/2}, \cdots, \boldsymbol{f}'_{n-1/2}, \boldsymbol{f}'_{n+1/2}, \cdots, \boldsymbol{f}'_{n+(K+1)/2}\right] \otimes \boldsymbol{I}_M \,.$$

Hence, given the CUT $\boldsymbol{x}$, the EFA gives N outputs:

$$\tilde{\boldsymbol{x}}_n = \boldsymbol{T}_n \boldsymbol{x} \,, \quad \text{for } n = 0, \ldots, N-1 \,.$$

The EFA reduced covariance matrices can be written as

$$\tilde{\boldsymbol{R}}_n = E\{\tilde{\boldsymbol{x}}_n \tilde{\boldsymbol{x}}_n^H\} = E\{\boldsymbol{T}_n \boldsymbol{x}\boldsymbol{x}^H \boldsymbol{T}_n^H\} = \boldsymbol{T}_n \boldsymbol{R} \boldsymbol{T}_n^H \,, \quad \text{for } n = 0, \ldots, N-1 \,,$$

which are used to calculate the $KN \times 1$ reduced-dimension Wiener-Hopf equations:

$$\tilde{\boldsymbol{w}}_n = \tilde{\boldsymbol{R}}_n^{-1} \tilde{\boldsymbol{s}}_n = (\boldsymbol{T}_n \boldsymbol{R} \boldsymbol{T}_n^H)(\boldsymbol{T}_n^H \boldsymbol{s}),$$

Although N output calculations are required for each received data vector, the transformed covariance matrices have dimensions $KN \times KN$, and hence are computationally favorable in most cases. For our simulations, $M = 6$ elements and $N = 32$, which requires inversions of 192×192 matrices when operating in the full-dimension space. Setting $K = 3$, the reduced-dimension matrices are 18×18. Recalling that matrix inversions are $O(P^3)$ for a $P \times P$ matrix, we see that $(MN)^3 \gg N(KM)^3$, which verifies the efficiency of the EFA.

2.5.2. Rank Modifiers

As an alternative to reducing the dimension of the problem, rank modifying techniques directly manipulate the covariance matrix estimate to increase its rank. The diagonal loading technique builds on the concept of retaining the principle components of the matrix and removing components with the smallest eigenvalues. To maintain a proper matrix rank, a scaled identity matrix $\hat{\sigma}_n^2 \boldsymbol{I}_{MN}$ is added, where $\hat{\sigma}_n^2$ is the background noise level estimate [32]. Essentially, the noise-only components are replaced with an average of the background, producing a better estimate.

The colored loading technique generalizes the diagonal loading technique by adding a quiescent covariance structure $\boldsymbol{R}_q$ to the estimate [33]:

$$\widehat{\boldsymbol{R}} \mapsto c_1 \widehat{\boldsymbol{R}} + c_2 \boldsymbol{R}_q \, ,$$

where $c_1, c_2 \geq 0$ and $c_1 + c_2 = 1$.The quiescent matrix *could be* an identity matrix as in the diagonal loading technique, but it could represent clutter structure known a priori. This includes previous estimates of the clutter ridge, possibly derived from known terrain information. By incorporating *a priori* knowledge, we can better suppress false alarms caused by sidelobes.

Covariance matrix tapers (CMTs) present another method of covariance manipulation, via element-wise multiplication with a positive semi-definite matrix [34]. Let $\boldsymbol{R}_{CMT}$ be the covariance matrix taper. Then the refined covariance matrix is given as

$$\widehat{\boldsymbol{R}} \mapsto \widehat{\boldsymbol{R}} \odot \boldsymbol{R}_{CMT} \, .$$

By constraining the second matrix to be positive semi-definite, the resultant matrix is positive definite, hence retaining the properties of a covariance matrix.

We can perform diagonal loading via CMT by applying

$$\boldsymbol{T}_{DL} = \mathbf{1}_{MN}\mathbf{1}_{MN}^{H} + \text{diag}\left\{\frac{\sigma^2}{\widehat{\boldsymbol{R}}_{1,1}}, \frac{\sigma^2}{\widehat{\boldsymbol{R}}_{2,2}}, \dots, \frac{\sigma^2}{\widehat{\boldsymbol{R}}_{MN,MN}}\right\},$$

where σ^2 is the scale factor, $\widehat{\boldsymbol{R}}_{m,n}$ represents the $(m,n)^{th}$ element of $\widehat{\boldsymbol{R}}$, and $\text{diag}\{\cdot\}$ denotes an operation that forms a diagonal matrix with its arguments on the diagonals. Color loading can be similarly performed.

Aside from increasing the rank of the matrix, CMTs are able to model internal clutter motion (ICM), which are pulse-to-pulse fluctuations of the received data. ICM effectively spreads the clutter ridge, as it is essentially an increase in temporal correlation. Different ICM models are discussed in Section 2.6.4.

2.5.3. Data Selection

Data selection algorithms indirectly modify the covariance matrix estimate by operating on the training data. A subset of the data is extracted that better homogenizes the training set. In particular, when we assume homogeneous clutter, we can filter range bins containing anomalous spikes ahead of time, since they will degrade the covariance estimate. This is the basis for the nonhomogeneity detector (NHD) algorithm, which compares the statistics of each range bin with the covariance matrix formed with the remaining training data to find

and remove anomalies [35]. That is, the statistics of the k^{th} range sample $\boldsymbol{x}_k$ is compared with the estimate

$$\widehat{\boldsymbol{R}} = \frac{1}{L-1}\sum_{l=1,l\neq k}^{L} \boldsymbol{x}_l \boldsymbol{x}_l^H,$$

where we assume we have L available range bins.

Power selected training (PST) takes a polar approach and opts to retain only the strongest range bins to overnull the data [36]. This may be desirable if the emphasis of the system is on the suppression of false alarms. The phase and power selected training is based on the PST, but also uses the phase information across the receiver channels to detect moving targets [37]. By omitting bins that contain moving targets, there is less chance of masking the targets of interest. The data selection methods are limited by the actual amount of homogeneous content available within the training set. If the clutter is too heterogeneous, we are left with a rather limited training set. This implies that the algorithm will likely be coupled with the aforementioned dimension-reducing or rank modifying techniques.

Overall, the methods discussed thus far attempt to mold the problem into that of detection in homogeneous clutter. This is conceptually different than deriving a method that is designed to operate in any heterogeneous situation, as in the knowledge-aided parametric covariance estimate (KAPE) approach. KAPE is a model-based approach that parameterizes the clutter ridge of the covariance matrix [38],[39]. For a given range R_o, we can describe the received clutter as the sum of the returns along N_p clutter patches:

$$\sum_{k=1}^{N_p} \alpha_k \boldsymbol{s}_k,$$

where α_k is the complex gain from the k^{th} clutter patch and $\boldsymbol{s}_k$ is the corresponding space-time steering vector. We see that the sum represents the returns along the clutter ridge and that we only need estimates of the complex gains. This is an over-simplified model, and in practice KAPE takes additional steps to ensure a realistic clutter model. Remarkably, KAPE is able to estimate the covariance structure by solely using the CUT, that is, it does not require a training set. This property makes it ideal to combat heterogeneous clutter, where the statistics may drastically vary between range bins. The multistep implementation of KAPE is the focus of the next section.

2.6. KAPE

The knowledge-aided parametric covariance estimation (KAPE) is a model-based approach to estimating the covariance matrix [38],[39]. KAPE is a knowledge-aided technique, i.e., a class of STAP algorithms that employ databases and a priori information to boost performance [40]. This includes any information about the array manifold, such as heading and known complex array gain errors, that affect the basic equation give in Section 2.1. When available, terrain information can also be used to accommodate null regions obscured by structures and to anticipate discretes in the data, such as moving vehicles.

To parameterize the clutter ridge of the scene, the iso-range ring of interest is divided into N_p constant-sized azimuth patches, were N_p is chosen to be smaller than the Doppler and spatial beamforming resolution. For this exposition, we assume a fixed elevation angle and negligible contributions from range ambiguities. To reiterate, the return from the CUT at range R_o can be written as

$$\sum_{k=1}^{N_p} \alpha_k \boldsymbol{s}_k$$

Let $\boldsymbol{S} = [\boldsymbol{s}_1, \dots, \boldsymbol{s}_{N_p}]$ and $\boldsymbol{P} = \text{diag}\left\{|E\{\alpha_1\}|^2, \dots, |E\{\alpha_{N_p}\}|^2\right\}$, where $\text{diag}\{\cdot\}$ denotes the formation of a diagonal matrix with the given arguments and $E\{\cdot\}$ denote the expectation of its argument. The covariance matrix is compactly written as

$$E[\boldsymbol{x}\boldsymbol{x}^H] = \boldsymbol{S}\boldsymbol{P}\boldsymbol{S}^H .$$

This formation assumes that the returns can be described with ideal steering vectors. We are able to estimate the gains $\alpha_1, \dots, \alpha_{N_p}$ using a single range bin, namely the CUT, by averaging over the redundancies of the pulse returns. Let $\boldsymbol{x}_n$ be the voltage corresponding to the n^{th} pulse and $\boldsymbol{s}_{k,n}$ be the $M \times 1$ ideal spatial steering vector corresponding to the k^{th} clutter patch and n^{th} pulse. We estimate the gain as

$$\alpha_k = \frac{1}{N}\sum_{n=1}^{N} \left\| \boldsymbol{s}_{k,n}^H \boldsymbol{x}_n \right\|^2 ,$$

the mean gain over the available pulses. If available, additional range bins may be used to further improve the estimate. For our purpose, we will use only the CUT to demonstrate KAPE's efficacy in highly heterogeneous clutter.

2.6.1. Array Errors

Realistically, the ideal space-time steering vectors $\boldsymbol{s_1}, \ldots, \boldsymbol{s_{N_p}}$ do not properly describe the returns as they do not account for inconsistencies such as sensor jittering and variances in the analog-to-digital voltage conversion, as well as complex gain errors in the spatial channels. We can divide the array errors into two cases: errors that are dependent on the steering direction $\boldsymbol{\epsilon_1}$ and angle-independent errors $\boldsymbol{\epsilon_o}$. By including the angle-independent errors, the scene is better modeled with the set of steering vectors

$$\boldsymbol{\epsilon_o} \odot \boldsymbol{s_{k,n}}\,, \qquad k = 1, \ldots, N_p, n = 1, \ldots, N\,.$$

We are less concerned with the angle-dependent errors, as $\boldsymbol{\epsilon_o}$ contributes most to the errors across the mainbeam [39]. We can utilize the given array information to obtain initial estimates of the errors $\hat{\boldsymbol{\epsilon}}_o$. This leads to a set of approximate spatial steering vectors:

$$\hat{\boldsymbol{v}}_{k,n} = \hat{\boldsymbol{\epsilon}}_o \odot \boldsymbol{s_{k,n}}\,, \qquad k = 1, \ldots, N_p, n = 1, \ldots, N,$$

which are used to derive the corresponding space-time steering vectors $\hat{\boldsymbol{v}}_1, \ldots, \hat{\boldsymbol{v}}_{N_p}$. Letting $\hat{\boldsymbol{V}} = [\hat{\boldsymbol{v}}_1, \ldots, \hat{\boldsymbol{v}}_{N_p}]$, the covariance matrix estimate is then $\hat{\boldsymbol{V}}\boldsymbol{P}\hat{\boldsymbol{V}}^H$.

Since KAPE is a parametric method, steering vector mismatches can lead to heavy losses, which we demonstrate in Section 4.3. Refined array error estimates are a necessity for successful applications by means of array calibration techniques. We are interested in *in situ* techniques that will adapt to any given clutter scene. These techniques are collectively referred to as cal-on-clutter methods, which include two notable methods: the maximum eigenvector method and the adjacent channel pairing (ACP) method [39].

The maximum eigenvector method assumes strongest returns from broadside of the array. In the case of a sidelooking linear array, this corresponds to the spatial steering vector $\mathbf{1}_M$, i.e., a vectors of ones. Since the principle eigenvector of the array manifold points towards the strongest power source, it follows that the principle eigenvector will ideally be a scaled vector of ones. In the presence of array errors, the broadside is perturbed, resulting in $\boldsymbol{\epsilon}_o \odot \mathbf{1}_M$. Likewise, the principal eigenvector follows this deviation. The maximum eigenvector method follows this concept by performing an eigendecomposition over the spatial covariance matrix from the available samples. The extracted principal component $\boldsymbol{q}_1$ gives us the array error estimate:

$$\boldsymbol{q}_1 = \hat{\boldsymbol{\epsilon}}_o \odot \mathbf{1}_M \ .$$

The spatial covariance matrix $\widehat{\boldsymbol{R}}_s$ is estimated by averaging over the available pulses and range bins:

$$\boldsymbol{R}_s = \frac{1}{NL} \sum_{n=1}^{N} \sum_{l=1}^{L} \boldsymbol{x}_{n,l} \boldsymbol{x}_{n,l}^H \ ,$$

where $\boldsymbol{x}_{n,l}$ represents the $M \times 1$ vector of values received from the n^{th} pulse of the l^{th} range bin.

The presence of multiple dominant scatterers about the iso-ring may bias the maximum eigenvalue method, as it violates the assumption that the broadside is strictly the strongest source. In such cases, the EFA centered at the DC bin can be employed to isolate the contributions from ground clutter at the expense of a reduced sample size for covariance estimation. This is the essence of the maximum eigenvector, Doppler centroid method [39].

Specifically, let $\boldsymbol{T}_o$ be the EFA transformation matrix centered about the DC bin, as elaborated in Section 2.5.1. Then, the centered covariance estimate is calculated as

$$\widetilde{\boldsymbol{R}}_s = \frac{1}{L}\sum_{l=1}^{L} \boldsymbol{T}_o^H \boldsymbol{x}_l \boldsymbol{x}_l^H \boldsymbol{T}_o,$$

where the average is performed over the available range bins.

The adjacent channel pairing method (ACM) takes an alternative approach by relating the phase differences between the channels. The phase progression between adjacent elements of uniform linear array is [41]

$$\exp(j2\pi df/c),$$

where d is the elemental spacing, c is the speed of light, and f is the transmit frequency. Since clutter maintains zero velocity for sidelooking radar steered at broadside, there will be zero progression between the elements. The adjacent pairing method assumes this and attributes any deviations in phase to array errors. Let $z_{m-1,n,l}$ and $z_{m,n}$ be the complex data received on the $(m-1)^{th}$ and m^{th} channels of the n^{th} received pulse and l^{th} range bin. Then the relative array error between the channels is estimated as:

$$\hat{\epsilon}_{m,n,l} = \frac{|z_{m,n,l}|}{|z_{m-1,n,l}|} \exp\left\{j\angle\left(z_{m,n} z_{m-1,n}^*\right)\right\},$$

which accounts for gain as well as the phase errors. Using the first channel as the reference, the relative channel errors are calculated for $m = 2, \ldots M$. The error estimates are subsequently refined by averaging over the N pulses and L available range bins:

$$\hat{\epsilon}_m = \frac{1}{NL}\sum_{l=1}^{L}\sum_{n=1}^{N}\hat{\epsilon}_{m,n,l}\,,\ \text{for } m = 1, \dots, M.$$

If N is large enough, the error estimates may be obtained using solely the CUT. That is, this method is applicable when $L = 1$. To calculate the absolute channel errors from the first channel, the relative errors are stacked, resulting in the array error estimate vector

$$\hat{\boldsymbol{\epsilon}}_o = [1 \quad \hat{\epsilon}_2 \quad \hat{\epsilon}_2\hat{\epsilon}_3 \quad \cdots \quad (\hat{\epsilon}_2\hat{\epsilon}_3 \cdots \hat{\epsilon}_M)]^T.$$

Note that the first element is always 1, as it is the reference channel. The ACM provides a computationally simple yet effective way of estimating array errors. However, if the transmit pulses fail to maintain constant contact with the target during the CPI, the array error estimates will be biased. This describes the case of a short dwell return. Furthermore, estimation errors propagate across the channels because relative array errors are stacked to calculate the absolute channel errors. This implies that

$$E\{|\epsilon_M - \hat{\epsilon}_M|\} > E\{|\epsilon_m - \hat{\epsilon}_m|\},\ \text{for } m = 2, \dots, M-1,$$

where ϵ_m is the actual array error of the m^{th} channel and $E\{\cdot\}$ calculates the statistical expectation of its argument. In general, the ACM method will degrade as the number of channels in the array increases.

2.6.2. Discrete Matched Filter

When multiple range bins are used for estimation, KAPE preemptively filters the data for anomalous points to better adapt to the underlying scene. The CLEAN algorithm, a method

of removing point-like returns, is well suited for this purpose [42],[43]. CLEAN is an iterative technique that, in this context, initially passes the data through a standard AMF filter using the SMI covariance matrix estimate. Upon locating a range-Doppler bin with a strong return, the exact Doppler and azimuth parameters of the outlier are further refined by applying a second AMF filter with a higher resolution. The power of the outlier is estimated and the anomalous discrete is subsequently subtracted from the data. This process is iteratively applied until no further outliers are detected.

A common concern is in the choosing of an appropriate threshold to determine the presence of a discrete. If the threshold is too low, CLEAN will be too aggressive and remove salient clutter structures. Furthermore, the algorithm needs to distinguish between a stationary discrete along the clutter ridge and a moving discrete away from it. This requires a second threshold, introducing another source of estimation error.

2.6.3. Clutter Ridge Shifts

Shifts in the clutter ridge result from uncertainties in the exact mainbeam steering vector. This effect may be more profound when the system experiences yaw during the signal acquisition [39]. Though small, these shifts introduce errors into the covariance parameterization, ultimately affecting the performance of KAPE. To estimate the perturbations, the KAPE algorithm dithers the Doppler frequency over a set of fractional frequencies f_Δ, shifting the ridge up and down the angle-Doppler spectrum. The $MN \times 1$ dithering vector is formulated as

$$\boldsymbol{y}_{dith}(f_\Delta) = \begin{bmatrix} 1 \\ \exp\{j2\pi f_\Delta / f_{PRF}\} \\ \cdots \\ \exp\{j2\pi f_\Delta (N-1)/f_{PRF}\} \end{bmatrix},$$

which is reminiscent of the temporal steering vector described in Chapter 1. The dithering is applied to the covariance matrix estimate via the following CMT:

$$\boldsymbol{C}_{dith}(f_\Delta) = \boldsymbol{y}_{dith}(f_\Delta)\boldsymbol{y}_{dith}^H(f_\Delta) \otimes \mathbf{1}_M.$$

KAPE incorporates the dithering frequency that best whitens the received data. The AMF is an excellent metric for this purpose, as it is intuitive and computationally efficient. Specifically, we calculate

$$c(f_\Delta) = \frac{1}{L}\sum_{l=1}^{L} \frac{\left|\boldsymbol{v}_o^H\left[\widehat{\boldsymbol{R}} \odot \boldsymbol{C}_{dith}(f_\Delta)\right]^{-1}\boldsymbol{x}_l\right|^2}{\boldsymbol{v}_o^H\left[\widehat{\boldsymbol{R}} \odot \boldsymbol{C}_{dith}(f_\Delta)\right]^{-1}\boldsymbol{v}_o},$$

where $\widehat{\boldsymbol{R}}$ is the current covariance estimate and $\boldsymbol{v}_o$ is the space-time steering vector centered about the mainlobe. Since ground clutter is concentrated within the mainlobe, $\boldsymbol{v}_o$ is a logical choice for the whitening metric. The frequency shift f_Δ that corresponds to the smallest AMF output $c(f_\Delta)$ is incorporated into the KAPE, since it best whitens the received data. Note that the AMF calculations can be computed in reduced-dimension space via EFA to reduce the computational load.

2.6.4. Clutter Ridge Spread

When the received pulses exhibit temporal decorrelation, the clutter ridge of the angle-Doppler spectrum spreads laterally in Doppler. This is intuitive, considering that we are calculating the covariance function. This type of spreading can be attributed to internal clutter motion (ICM), which is induced by radar errors and clutter fluctuations in the environment.

Since KAPE sums only the bins along the diagonal, severe spectral leakage can lead to under-nulling bins adjacent to zero Doppler. To compensate, KAPE employs covariance matrix tapers (CMT) that effectively model ICM. For example, the Gaussian temporal autocorrelation CMT is given as [17],[39]

$$\boldsymbol{R}_{CMT} = \text{toeplitz}\left\{1, \exp(-\tau^2\rho), \exp(-(2\tau)^2\rho), \ldots, \exp\left(-\left((N-1)\tau\right)^2\rho\right)\right\},$$

where $\text{toeplitz}\{\cdot\}$ denotes an operation where a Toeplitz matrix is formed from its arguments. The equation includes the constant ρ that controls the degree of the clutter ridge spread. The Gaussian model has been demonstrated to match water clutter. In practice, several values of ρ must be tested to find an appropriate fit to the given CUT. Similar to the clutter shift matching process discussed in Section 2.6.3, the AMF is employed to find the CMT parameters that best whiten the received data.

Another popular CMT for modeling clutter ICM is the Billingsley model, which has demonstrated its efficacy in modeling land clutter spread [4],[14]. Its power distribution is given by

$$P(f) = \frac{r}{r+1}\delta(f) + \frac{1}{r+1}\frac{\beta}{2}\exp(-\beta|f|),$$

where $\delta(\cdot)$ is the Dirac delta function and β is the shape parameter. The first and second terms represent the stationary and moving components, respectively, where r is the power ratio between them.

2.7. Performance Metrics

This section discusses various metrics for comparing the performance of different algorithms. This includes receiver-operating characteristic (ROC) curves, SINR losses, and exceedance plots. Each metric offers a different perspective on the detection and classification problem.

2.7.1. ROC

A receiver-operating characteristic (ROC) curve is a plot depicting the possible combinations of probability of false alarms P_{FA} and probability of detections P_D obtainable using the classification algorithm. Every point on the ROC curve corresponds to a threshold T; a detection is considered a target if it exceeds T. Hence, false alarms correspond to the case when non-targets produce a detection value over T and correct detections correspond to the case when targets produce a detection value over T. The innate tradeoff is that, although we reduce false alarms by setting the threshold high, we also lose correct detections as well.

For our algorithm assessments, the beam is fixed at broadside. Detections are formed using the AMF on the Doppler frequency bins that are over the minimum detectable velocity (MDV). Targets below the MDV are masked by ground clutter and are assumed to be undetectable. Two sets of data are generated for this analysis: clutter-only data and clutter-plus-target data. The clutter-only data is initially generated using the cell-based clutter model (CCM) explained in Section 3.1 and the clutter-plus-target data is generated by injecting a Swerling 1 target of power σ_t^2 into the frequency bins of interest. The ROC curve is formed empirically by averaging over N_s trials.

2.7.2. Exceedance Plots

The exceedance plot is a plot of the P_{FA} versus the detection threshold T. Hence, no additional information is needed to generate these plots, given data for the ROC curves. However, the tails of the exceedance plot indicate the presence of unsuppressed spiky clutter. For example, strong clutter residue requires a larger threshold to mitigate, which translates to long tails on the plots. Overall, this allows us to analyze a system's ability to remove discrete-like clutter, such as CDs, TSDs, and residue from heterogeneous clutter. The exceedance curve obtained from homogeneous clutter gives us a benchmark for performance. If the STAP algorithm performs properly, the output clutter is whitened.

2.7.3. SINR Loss Plots

There are several signal-to-interference-plus-noise ratio (SINR) plots that are of interest. The standard SINR power for STAP weights $\boldsymbol{w}$ corresponding to the space-time steering vector $\boldsymbol{s}$ is

$$SINR_{opt} = \frac{\sigma_t^2 |\boldsymbol{w}^H \boldsymbol{v}|^2}{\boldsymbol{w}^H \boldsymbol{R} \boldsymbol{w}} = \sigma_t^2 \boldsymbol{v}^H \boldsymbol{R}^{-1} \boldsymbol{v} ,$$

where $\boldsymbol{R}$ is the clairvoyant null hypothesis covariance matrix and σ_t^2 is the target power. We refer to this as the optimal SINR, as it requires the knowledge of $\boldsymbol{R}$. The optimal SINR loss $L_{s,1}$ is the optimal SINR normalized by the SNR of the system:

$$L_{s,1} = \frac{SINR_{opt}}{SNR} = \frac{\sigma_t^2 \boldsymbol{v}^H \boldsymbol{R}^{-1} \boldsymbol{v}}{(\sigma_t^2/\sigma_n^2)MN} = \frac{\sigma_n^2}{MN} \boldsymbol{v}^H \boldsymbol{R}^{-1} \boldsymbol{v} ,$$

where σ_n^2 is the background noise power. Essentially, $L_{s,1}$ computes the losses incurred by the addition of ground clutter. In the absence of ground clutter, $\boldsymbol{R} = \sigma_n^2 \boldsymbol{I}_{MN}$ and hence $L_{s,1} = 1$, which is the upper bound of the performance metric. For analysis, we fix the space-time steering vector $\boldsymbol{v}$ to be steered towards 0^o elevation and 0^o azimuth while varying the Doppler frequency to span the frequency space.

For this work, we are most interested in the adaptive SINR loss, which is the loss incurred due to the STAP weights being calculated with the covariance matrix estimate $\widehat{\boldsymbol{R}}$. The adaptive SINR loss $L_{s,2}$ is calculated as

$$L_{s,2} = \frac{SINR_{adp}}{SINR_{opt}} = \frac{\sigma_t^2 |\widehat{\boldsymbol{w}}^H \boldsymbol{v}|^2}{\widehat{\boldsymbol{w}}^H \boldsymbol{R} \widehat{\boldsymbol{w}}} \Big/ SINR_{opt} = \frac{|\widehat{\boldsymbol{w}}^H \boldsymbol{v}|^2}{\widehat{\boldsymbol{w}}^H \boldsymbol{R} \widehat{\boldsymbol{w}}} \Big/ \boldsymbol{v}^H \boldsymbol{R}^{-1} \boldsymbol{v} \,,$$

where $\widehat{\boldsymbol{w}} = \widehat{\boldsymbol{R}}^{-1} \widehat{\boldsymbol{v}}$ with the approximate steering vector $\widehat{\boldsymbol{v}}$. In the absence of array errors and, assuming the received data vectors are independent and identically distributed (i.i.d.), the SMI method can achieve an average adaptive SINR loss of 3 dB when $L = 2MN$ data samples are available [7]. Commonly known as the Reed, Mallet, and Brennan (RMB) rule, this provides a benchmark for new algorithms. For example, 0 demonstrates that KAPE attains a better adaptive SINR loss without assuming i.i.d. data. In general, SMI is performed in a reduced-dimension space due to the limited sample size and to lessen the computational complexity. Let $\boldsymbol{T}_m$ be the EFA transformation matrix centered around the m^{th} Doppler bin. Then the corresponding reduced-dimension STAP weight vector is

$$\widetilde{\boldsymbol{w}}_m = \left(\boldsymbol{T}_m^H \widehat{\boldsymbol{R}} \boldsymbol{T}_m\right)^{-1} (\boldsymbol{T}_m^H \widehat{\boldsymbol{v}})$$

and the adaptive SINR is calculated as

$$SINR_{adp}(m) = \frac{|\widetilde{\boldsymbol{w}}_m^H (\boldsymbol{T}_m^H \boldsymbol{v})|^2}{\widetilde{\boldsymbol{w}}_m^H (\boldsymbol{T}_m^H \boldsymbol{R} \boldsymbol{T}_m) \widetilde{\boldsymbol{w}}_m} \,.$$

The optimal SINR is calculated similarly. Overall, the adaptive SINR loss $L_{s,2}$ shows us the losses incurred using STAP, where a value of 0 dB indicates no loss.

Chapter 3

DETECTION IN HETEROGENEOUS CLUTTER

Textured clutter refers to background clutter that can be described with an underlying distribution model, as described in Section 2.3. In the homogeneous case, the complex Gaussian distribution is assumed. We are interested in heterogeneous cases, i.e., cases when the Gaussian model no longer suffices.

Recall that the K-distribution is well suited for sea clutter observed by high-resolution radars, since it encapsulates cases of spiky states. The compound model representation of K-distributed random variable is

$$f(x) = \int_0^\infty f(x|\gamma) f(\gamma) d\gamma\,,\ x \geq 0,$$

where $f(x|\gamma)$ is the Rayleigh probability density function conditioned on γ and $f(\gamma)$ is the gamma density function. We have chosen this model for analysis, as it has been shown to provide good empirical fits to real data using only two free parameters. Hence, our objective is to find a STAP method suitable for K-distributed clutter. The first step in our analysis is to accurately simulate the heterogeneous clutter. We introduce our novel cell-based clutter model (CCM) for clutter generation, adopting the idea of compound distributions.

3.1. Cell-Based Clutter Model

The modulating Rayleigh random variable of the K-distribution is generated by taking the magnitude of a zero-mean, unit-variance, complex Gaussian random variable. This results in a unit-mean, unit-variance Rayleigh process [13]. Following the compound model, this leaves two parameters for estimation: the scale and shape of the gamma distribution. For our simulations, we adopt the gamma ratio $r_\gamma = E[\gamma]/\mathrm{std}[\gamma]$, which details the spikiness of the scene as explained in Section 2.3.1. The mean is kept at a fixed value for intuitive performance comparisons.

The scene of interest is divided into cells with sizes finer than the range resolution of the system, capable of delineating disparate clutter regions. A new simulation is initiated by assigning each cell a draw from the gamma distribution with a gamma ratio that aptly describes the clutter roughness. Recall from Section 2.1 that we model the returns as the sum of the contributions from the clutter patches:

$$\sum_{k=1}^{N_p} \alpha_k \boldsymbol{s}_k \, .$$

The patch reflectivity ρ_k corresponding to the k^{th} patch is calculated by averaging the intersecting scene cells as illustrated in Figure 3. The radar cross section (RCS) per area is determined using the constant gamma model [11],[12]:

$$\sigma^o = \rho_k \sin\psi \, ,$$

where ψ is the grazing angle. Finally, the k^{th} patch gain is calculated as

$$\alpha_k = g_k\sqrt{\sigma^o A}\,,$$

where A is the area of the patch and g_k is the beamformer voltage gain factor.

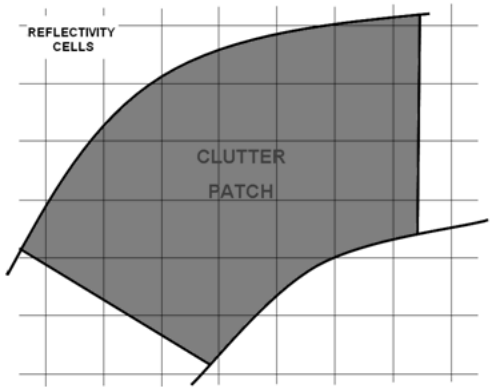


Figure 3. The clutter patch reflectivity is calculated by summing the individual cells of the CCM.

The resulting gain is also gamma distributed since it is the sum of multiple independent gamma random variables with a fixed scale parameter [13]. Hence, the resulting compound distribution is K-distributed.

The CCM can also generate homogeneous clutter by assigning a constant value γ_o to each cell, which implies that all patch reflectivity values $\rho_1, \rho_2, \ldots, \rho_{N_p}$ will be equivalent. Referring to the equation for a compound distribution, this reduces the model to a multiplication between a Rayleigh distribution and a constant. The resulting distribution is a scaled Rayleigh, which describes a homogeneous scene.

The CCM affords the flexibility to simulate radar returns from various radar systems without having to re-estimate the gamma parameters. This is because the gamma parameters describe the clutter scene itself, detached from the radar system. Hence, e.g., we may vary the range

resolution, pulse frequencies, and the number of array elements for comparative analysis over a common scene of interest. If available, terrain databases may be utilized to estimate the required gamma parameters, *a priori*.

Since the clutter patch size is large for low-resolution radar systems, each patch will intersect with more scene cells of the CCM. This implies that the RCS will be an average of many independent and identically distributed random variables with common distributions. By the central limit theorem, the result will converge to the Gaussian distribution. This is correct in a phenomenological sense, since returns from low-resolution radar systems are typically described with the homogeneous clutter model, and hence Gaussian returns. Conversely, the clutter patches will intersect with few scene cells in high-resolution systems, hence maintaining the shape of the original gamma distribution. When the scene is generated with a low gamma ratio, corresponding to spiky texture, the radar return will maintain its heterogeneity. This also matches with what is phenomenologically observed. Hence, the CCM is a simple and intuitive approach to generating clutter that produces returns congruent with real clutter.

3.2. KAPE in Heterogeneous Texture Clutter

We can now use the CCM model we developed to reliably simulate various samples of heterogeneous clutter. The performance of STAP algorithms in spiky scenes should reflect desired real world capacity. We first demonstrate losses incurred using SMI in heterogeneous texture clutter. Recall that SMI is tailored for homogeneous backgrounds, and does not adapt well to spikes in the data. In contrast, we implement the KAPE algorithm, which has

the ability to adapt solely using information from the CUT. The nulls of the STAP filter do not average out as in the SMI case.

For the simulations, the MDV is set to include Doppler frequencies with SINR losses above -15 dB. We inject 0 dBsm Swerling 1 targets into the clutter-only data to generate clutter-plus-target data used to form the ROC curves. For the SMI method, the covariance matrix estimate is obtained by averaging the available data samples within the range swath. In contrast, solely using the CUT, KAPE parameterizes the covariance matrix estimate using the steering vectors along the clutter ridge with estimated powers. For both the SMI and KAPE methods, we apply the EFA algorithm on the simulated receive data for dimension reduction, followed by the AMF to obtain the detection statistic. The detection statistics are compared to an array of fixed thresholds T to calculate the performance metrics.

Homogeneous clutter was generated using the CCM by setting the summed reflectivity values to a constant γ. The constant γ was chosen such that the single-channel, single-pulse clutter-to-noise ratio (CNR) power is 25 dB. Heterogeneous K-distributed clutter was also generated with the CCM, but with a gamma ratio of 0.1. Recall from Figure 2 that this corresponds to a spiky scene. The mean of the gamma distribution was set to maintain an average CNR of 25 dB, as in the homogeneous case. We generate $L = 36$ range samples per simulation run, including the CUT. Note that $L < MN$ in our case, which calls for a dimension reducing algorithm, namely the EFA.

The radar system we are simulating is a sidelooking airborne radar employing a uniform linear array with $M = 6$ elements. It transmits $N = 32$ pulses per coherent processing interval (CPI). For the EFA algorithm, we compute a filter bank of outputs with each filter

processing $K = 3$ adjacent Doppler bins. The key parameters of the simulation are listed in Table 2.

Table 2. Radar systems parameters used to generate the received data.

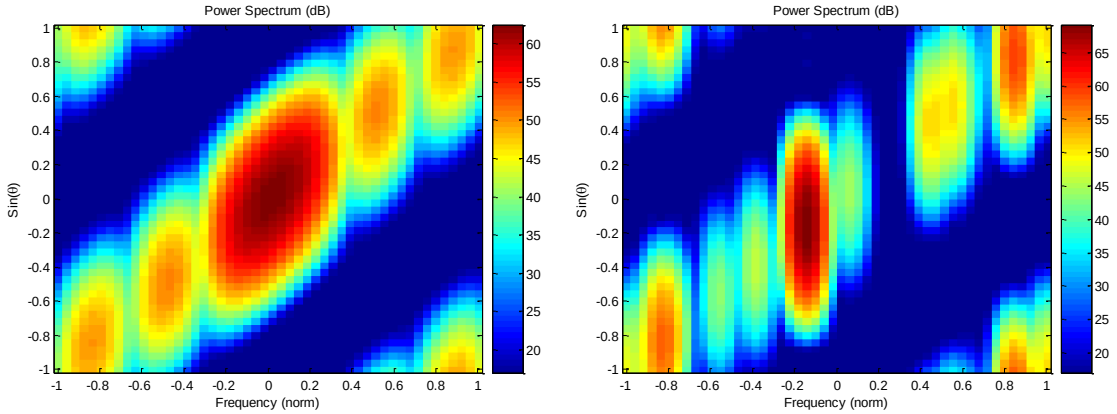
M = 6 channels	Radar Frequency = 1 GHz
N = 32 pulses per CPI	PRF = 750 Hz
K = 3 adjacent Doppler bins	Radar Resolution = 5 m
Swath Center = 30 km	Mean CNR = 25 dB
Aircraft Velocity = 110 m/s	Number of Range Samples = 2KM = 36

To better describe the problems with heterogeneous texture clutter that we wish to address, this section presents some power spectrum density maps. To calculate the power response of the covariance matrix $\boldsymbol{R}$, we form the space-time steering vectors $\boldsymbol{v}(f_i, \phi_j)$, for $i = 0, \dots, I-1$, $j = 0, \dots, J-1$ that uniformly spans the Doppler and azimuth space. The power is calculated as

$$P(i,j) = \left[\boldsymbol{v}(f_i, \phi_j) \odot \boldsymbol{w}_{hamm}\right]^H \boldsymbol{R}\left[\boldsymbol{v}(f_i, \phi_j) \odot \boldsymbol{w}_{hamm}\right],$$

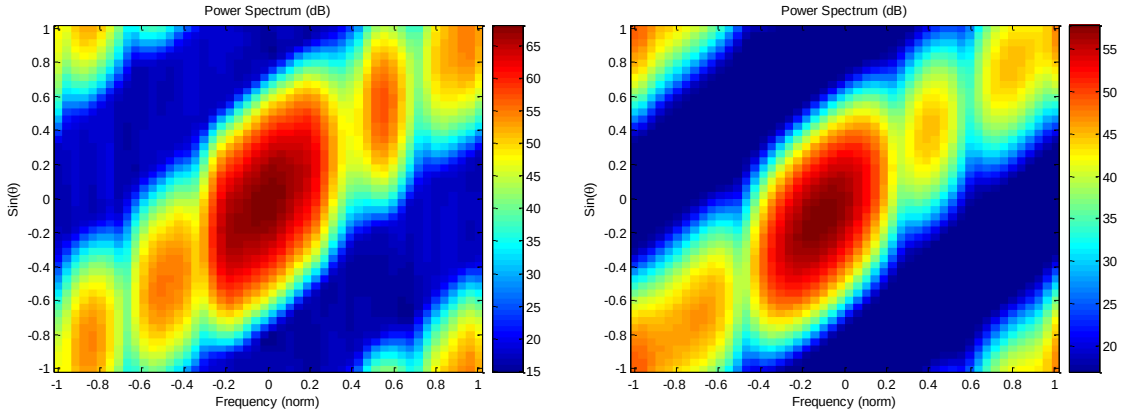
where $\boldsymbol{w}_{hamm}$ is the Hamming window used to suppress sidelobes. Figure 4(a) illustrates the power spectrum density (PSD) map for homogeneous clutter using the parameters of Table 2. Note that the low resolution along the azimuth is attributed to the small number of channels in the system. Figure 4(b) illustrates the PSD map for K-distributed clutter with a gamma ratio of 0.1. The volatile scene produces a strong reflection centered around the Doppler bin corresponding to -0.2. Hz that overpowers the previous ground clutter peak centered about 0^o azimuth angle and 0 Hz Doppler frequency.

Figure 5(a) is the PSD estimate obtained using the SMI to estimate $\widehat{\boldsymbol{R}}$. We see that the peak point is still centered around 0 Hz Doppler and 0^o azimuth because it is the average of the received data vectors. In contrast, Figure 5(b) illustrates the PSD obtained with KAPE, which is clearly centered about the dominant clutter of the heterogeneous scene. Since KAPE better matches the CUT, we can expect improved performance.



(a) (b)

Figure 4. Example power spectrum density maps of clutter. (a) PSD of homogeneous clutter. (b). PSD for heterogeneous texture clutter with a gamma ratio of 0.1.



(a) (b)

Figure 5. Example power spectrum density maps in the presence of heterogeneous texture clutter. (a) Power spectrum density estimate obtained using the SMI method. (b). Power spectrum density estimate obtained via KAPE.

Figure 6 compares the ROC curves of the simulations corresponding to several cases of interest. The solid line demonstrates that the SMI performs well in homogeneous clutter, as expected. For example, we approximately attain a 0.84 probability of detection P_D for a system that maintains a 5×10^{-4} probability of false alarm P_{FA}. However, severe degradations are incurred in the presence of K-distributed clutter, shown as the dash-dot plot. Since the SMI averages over the entire range interval, the nulls for spiky range cells are insufficiently deep. For the same system that maintains $P_{FA} = 5 \times 10^{-4}$, the probability of detection diminishes to approximately 0.64, which is unacceptable in practical cases of interest.

This effect of K-distributed clutter is further demonstrated in the exceedance plots of Figure 7. The heavy tail is the result of strong clutter residues that the STAP filter did not adequately suppress. Specifically, to maintain $P_{FA} = 5 \times 10^{-4}$ in heterogeneous clutter, the threshold is increased from 35.21 to 92.72, a difference of 57.51. Figure 8 plots the P_D obtained as a function of the fixed threshold used for detection. In the case of SMI, the increase in the threshold from 35.21 to 92.72 reduces the P_D to 0.64, as expected. The red dotted lines of Figure 8 delineate the results of this case.

In contrast, KAPE performs well in K-distributed clutter, as demonstrated by the dashed curve of Figure 6, hence demonstrating its effectiveness in suppressing spiky clutter. Comparing with our previous examples, KAPE attains a P_D of 0.92 for a system that maintains a 5×10^{-4} probability of false alarm P_{FA}. Note that this is better than what is attainable with the SMI in homogeneous clutter. Figure 9 plots KAPE ROC curves in

homogeneous and heterogeneous clutter to demonstrate minimal losses incurred between the cases.

Figure 7 shows the exceedance plot of KAPE in heterogeneous clutter. The short tails indicate that KAPE is able to generate the deep nulls required to suppress clutter residue. In fact, the tails are very similar to those found using SMI in homogeneous clutter, which explains the minimal P_D losses observed in Figure 9. Note that the KAPE covariance matrix estimates generally have higher power levels than those found with the SMI method, which accounts for the offset seen in the exceedance plot of Figure 7. Since this is just a scaling factor, it does not affect the performance. Overall, KAPE outperforms SMI in homogeneous clutter since it is a model-based parametric technique, unaffected by the SINR loss due to averaging.

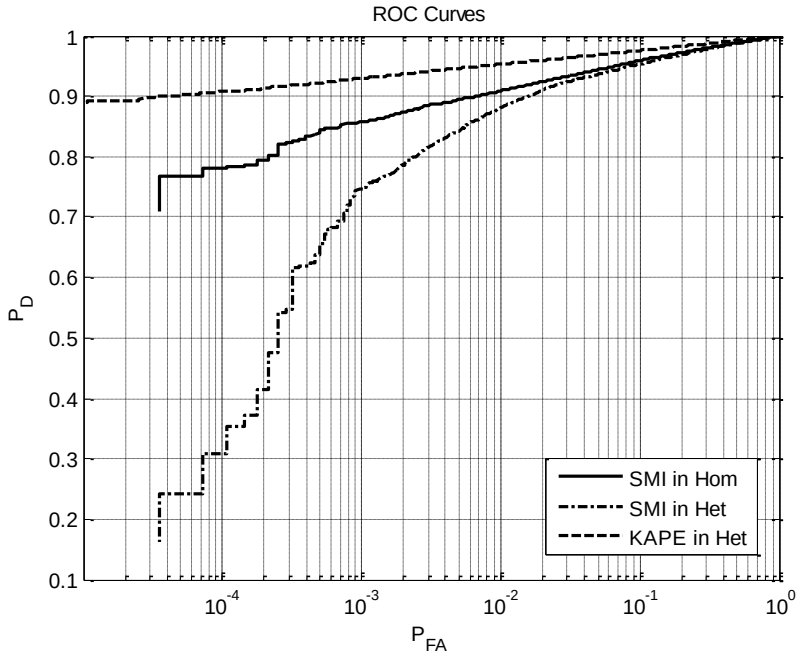


Figure 6. ROC curves comparing performances of KAPE and SMI in heterogeneous texture clutter.

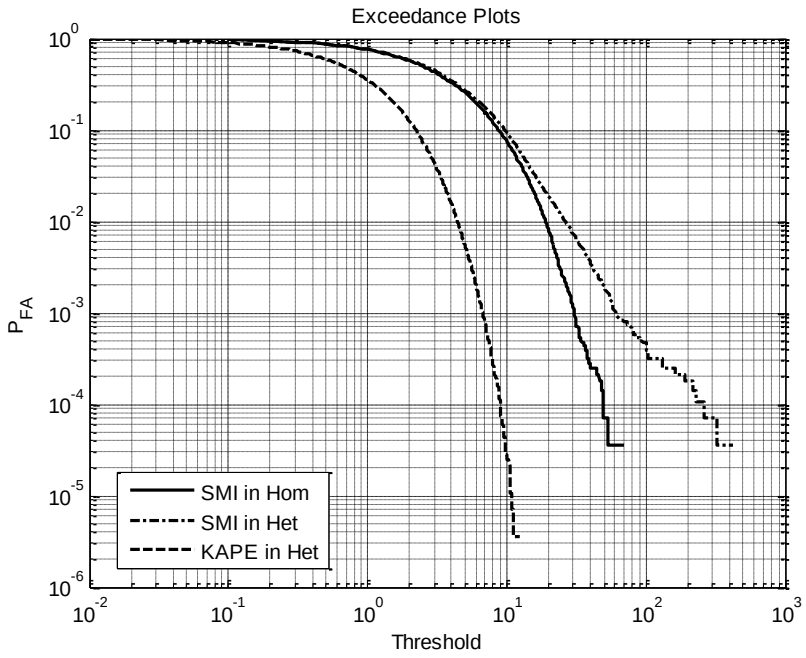


Figure 7. Exceedance plots demonstrating the clutter residue between the KAPE and SMI algorithms.

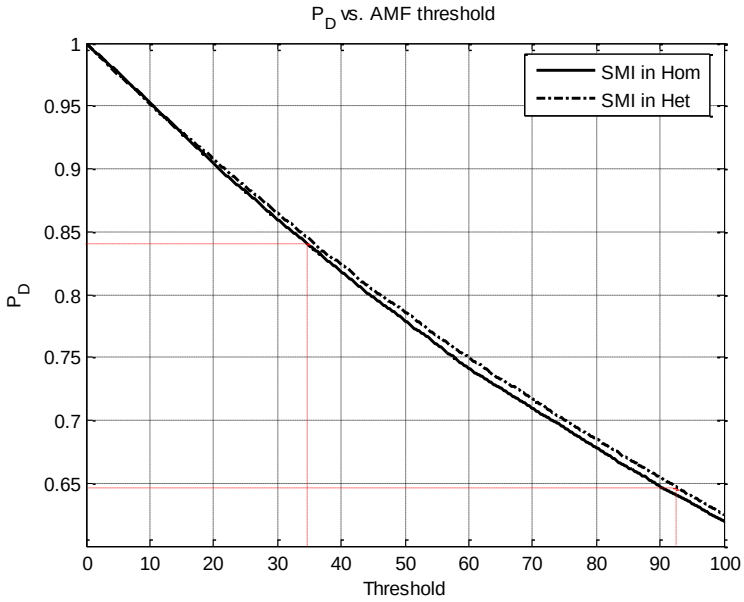


Figure 8. Plot of attainable probability of detection using SMI in homogeneous and heterogeneous clutter.

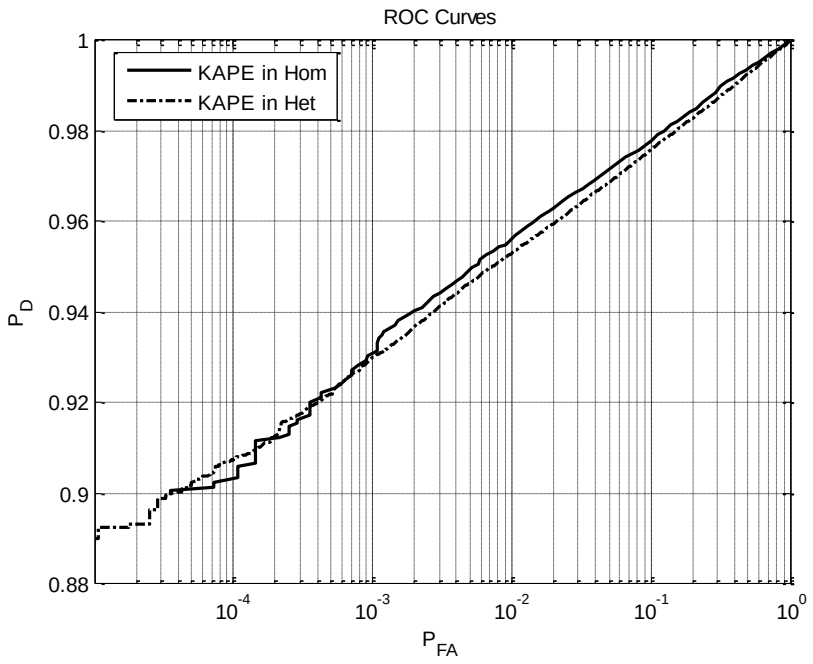


Figure 9. ROC curves demonstrating minimal KAPE performance loss in the presence of heterogeneous clutter.

Chapter 4

DETECTION IN THE PRESENCE OF DISCRETE INTERFERENCE

In addition to its capabilities in clutter with heterogeneous clutter, KAPE has been shown to perform well in the presence of clutter discretes (CD) and targets in secondary data (TSD) [22]. The ability to adapt over a single range bin allows KAPE to create deeper nulls to cancel spurious clutter, while avoiding signal cancellation due to off-ridge nulling.

4.1. KAPE Limitations

The computational requirements of KAPE are an inherent limitation in real applications. An initial step in KAPE requires an estimate of the gains from the N_p azimuth patches, once for each range bin. To ensure a good model fit, N_p is chosen to be finer than the achievable Doppler and beamforming resolution. The overall number of estimates becomes nontrivial. However, the biggest burden is the calculation of the inverse covariance matrix $\widehat{\boldsymbol{R}}^{-1}$ required in the formulation of the STAP weights. Even for a modest system with $M = 6$ and $N = 32$, such as the one analyzed in Section 3.2, this equates to 192×192 matrix inversions for all range bins. This calculation may be infeasible for real-time systems, hence prohibiting the implementation of KAPE. In the next subsection, we introduce a new technique we call the Gram Schmidt (GS) KAPE, which introduces computationally efficient modifications to

KAPE covariance estimation. GS KAPE circumvents the need for direct matrix inversions by formulating a parametric model for estimating $\widehat{\boldsymbol{R}}^{-1}$.

Another limitation of KAPE is its sensitivity to model errors. Although array errors adversely affect SMI, its impact on KAPE can be severely limiting, since array errors translate to incorrect steering vectors in the covariance estimate equation, implying a model mismatch. Hence, array calibration techniques are needed in the implementation of KAPE. Section 4.3 showcases the improvements attained from known calibration methods. We then introduce our power-ridge alignment (PRA) method of array calibration, which outperformed the standard methods we tested against.

4.2. Gram-Schmidt KAPE

We have developed our Gram-Schmidt (GS) KAPE method as a means to reduce the computational burden of the original KAPE algorithm. Given a set of vectors that span $\mathbb{R}^n$, the standard Gram-Schmidt (GS) algorithm iteratively derives a set of orthogonal vectors that span the same space [44]. For our purpose, we would like to orthogonalize the set of steering vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{N_p}$ that span the clutter ridge, via the GS algorithm, to form the orthonormal set $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_q$. We see that $q \leq N_p$ since the original set of steering vectors oversample the clutter space; i.e., the rank of $\boldsymbol{V}$ is less than N_p. However, since N_p is typically much larger than q, the elements of $\boldsymbol{V}$ exhibit a high degree of linear dependency. This precludes a direct application of GS, because it produces unstable results due to the limits of numerical precision.

To illustrate this problem, recall that the usual GS procedure obtains the k^{th} orthonormal vector by computing:

$$\boldsymbol{u}_k' = \boldsymbol{v}_k - \sum_{i=1}^{k-1} (\boldsymbol{v}_k^H \boldsymbol{u}_i)\boldsymbol{u}_i,$$

$$\boldsymbol{u}_k = \boldsymbol{u}_k' / \|\boldsymbol{u}_k'\| \, .$$

Since the inputs are presented in an arbitrary order, $\boldsymbol{v}_k$ may be approximately linearly related to the previous vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1}$. In that case, we see that $\|\boldsymbol{u}_k'\| \cong 0$, producing unstable results. Furthermore, we obtain $\boldsymbol{u}_k$ that mostly span the non-clutter space.

In light of the aforementioned problems, we implemented a modified version of GS that judiciously orders the input vectors to circumvent numerical instabilities. This requires an extra variable, namely a threshold, that defines the stopping criterion of the algorithm and determines the number of output vectors q. This presents an inherent tradeoff: we desire a small q for computational efficiency, but it also needs to be large enough to sufficiently span the clutter space. Our experiments showed that if q is too large, we inadvertently include orthogonal vectors that do not span the clutter space, hence only increasing unwanted background noise in the covariance estimate.

To initialize the modified GS, the first input is chosen arbitrarily as the first vector in the list, i.e.,

$$\boldsymbol{v}_{(1)} = \boldsymbol{v}_1,$$

where $\boldsymbol{v}_{(j)}$ indicates the j^{th} sorted input vector. We form $\boldsymbol{u}_1$ by normalizing $\boldsymbol{v}_{(1)}$ and enumerate a list of residue vectors from the remaining inputs:

$$\boldsymbol{r}_j = \boldsymbol{s}_j - \boldsymbol{s}_j^H \boldsymbol{u}_1 \,, \quad \text{for } j = 2, \dots, N_p \,.$$

The index corresponding to the largest residue is then located:

$$j_{\max} = \text{argmax}_j \left\| \boldsymbol{r}_j \right\| .$$

Hence, the second input is the input vector corresponding to the max index $\boldsymbol{v}_{(2)} = \boldsymbol{v}_{j_{\max}}$, as it presents the most information on the clutter space. This process is iterated to obtain additional othonormal vectors.

At the start of the i^{th} iteration, we have the orthonormal set $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{i-1}$ from the sorted input vectors $\boldsymbol{s}_{(1)}, \dots, \boldsymbol{s}_{(i-1)}$. As before, we enumerate a list of residue vectors using the input vectors and all output orthonormal vectors:

$$\boldsymbol{r}_j = \boldsymbol{s}_j - \sum_{k=1}^{i-1} \boldsymbol{s}_j^H \boldsymbol{u}_k \,, \quad \text{for } j = 1, \dots, N_p \,.$$

Note that in the actual implementation, the cumulative summation can be done iteratively, hence reducing the computational load. Also, the residues are only calculated for the unused input vectors, although we write it for all $j = 1, \dots, N_p$ for the simplicity of exposition. As before, we locate the index corresponding to the largest residue:

$$j_{\max} = \text{argmax}_j \left\| \boldsymbol{r}_j \right\| .$$

At this point, we verify whether the best candidate input exceeds the predefined threshold value T_{GS}:

$$\left\|\boldsymbol{r}_{j_{\max}}\right\| \geq T_{GS} .$$

If the criteria is met, we continue the iterative procedure with $i = j_{\max}$. Otherwise, we stop, and declare the number of output vectors as $q = i - 1$ with the orthonormal output set $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{i-1}$. Judiciously choosing the inputs avoid the problems associated with presenting the GS algorithm with completely redundant vectors.

After obtaining the orthonormal vectors $\boldsymbol{U} = [\,\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \cdots \quad \boldsymbol{u}_q]$, we rewrite the original KAPE covariance matrix equation:

$$\sum_{k=1}^{q} \beta_k \boldsymbol{u}_k \boldsymbol{u}_k^H,$$

where we have substituted the original steering vector set $\boldsymbol{V}$ with $\boldsymbol{U}$ and its associated power estimates $\boldsymbol{B} = \operatorname{diag}\left\{|\beta_1|^2, \ldots, \left|\beta_q\right|^2\right\}$. In this case, only q gains need to be estimated, which is an improvement over N_p and an overall reduction in computations. The estimates are obtained using the method explained in Section 2.6. In matrix form, the total interference covariance matrix is written as

$$\widehat{\boldsymbol{R}} = \boldsymbol{U}\boldsymbol{B}\boldsymbol{U}^H + \sigma_n^2 \boldsymbol{I}_{MN} ,$$

including the white Gaussian background noise of power σ_n^2. The biggest change from the original KAPE formulation is the orthonormality of the column vectors of $\boldsymbol{U}$, enabling us to parameterize the inverse covariance matrix as well:

$$\widehat{\boldsymbol{R}}^{-1} = \boldsymbol{U}\boldsymbol{B}_{TOT}^{-1}\boldsymbol{U}^H + \frac{1}{\sigma_n^2}(\boldsymbol{I}_{MN} - \boldsymbol{U}\boldsymbol{U}^H)\,.$$

We introduce the term

$$\boldsymbol{B}_{TOT}^{-1} = \mathrm{diag}\left\{\frac{1}{\beta_1^2 + \sigma_n^2}, \ldots, \frac{1}{\beta_q^2 + \sigma_n^2}\right\},$$

which combines the gain estimates and the background noise.

This formulation circumvents the need to calculate direct inverses of $MN \times MN$ matrices required in the standard KAPE algorithm. Typically, to bypass this complication, KAPE applies dimension reducing algorithms such as EFA to divide the inversion into smaller problems. Since this step is unnecessary in GS KAPE, we additionally avert the SINR loss associated with dimension-reduction.

To demonstrate the efficacy of GS-KAPE with the lessened computational load, we compare it to the standard KAPE algorithm of Section 2.6. The performance is averaged over 1000 trials of simulated homogeneous clutter data with single-channel, single-pulse clutter power of 25 dB. The system receives data over a uniform linear array with $M = 6$ channels and $N = 32$ pulses. We have access to $L = 36$ range bins to estimate the clutter covariance matrix. The standard KAPE algorithm requires gain estimates over $N_p = 656$ clutter patches, whereas the GS-KAPE condenses this amount to just $q = 37$ orthogonal

patches. This is over 17 times in computational savings. For a fair comparison, both the standard KAPE and GS-KAPE are filtered with the EFA algorithm to reduce the dimension. Note that GS-KAPE does not actually require this step, since we have a parametric equation for the inverse covariance matrix. Figure 10 compares the SINR curves between the methods, along with the SMI case. The SMI case suffers SINR loss associated with averaging over the range bins, where the degree of loss is a function of the number of available homogeneous range samples available. Figure 11 shows the ROC curve for the same cases, where a 0 dB Swerling 1 target was injected to compute the probabilities of detection. The results were averaged over 1000 trials. Overall, we see that GS KAPE suffers no significant performance loss.

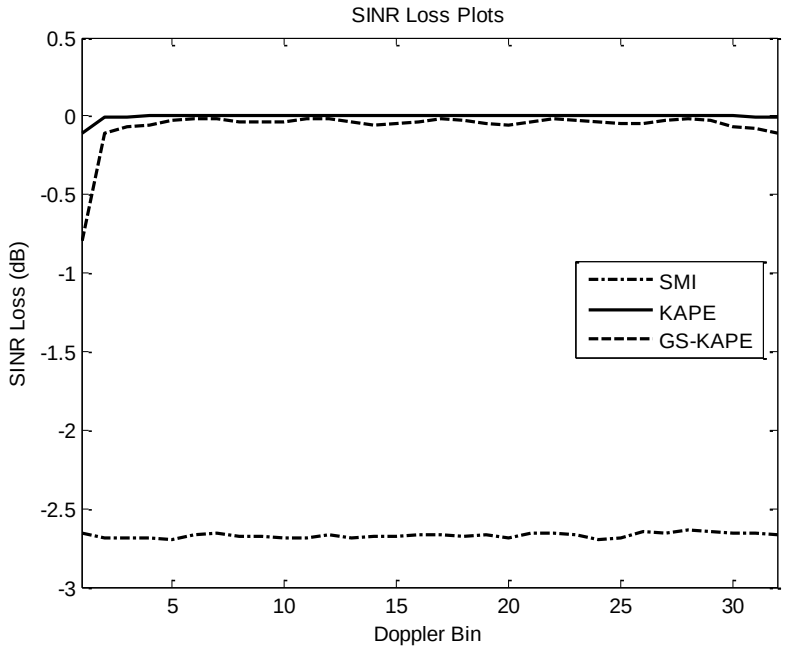


Figure 10. SINR loss plots demonstrating the efficacy of GS KAPE.

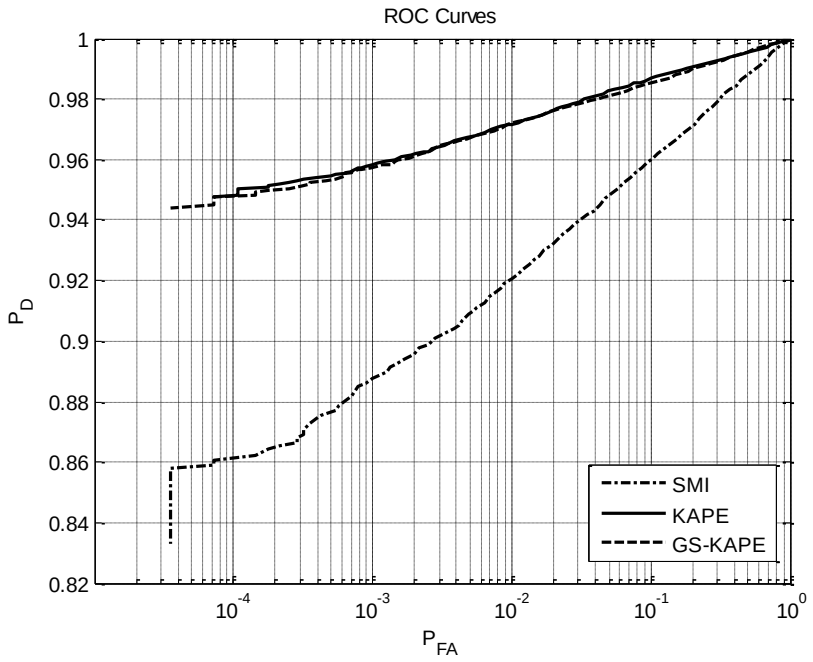


Figure 11. ROC curves demonstrating the efficacy of GS KAPE.

Several alternative methods can be employed to obtain orthogonalized steering vectors, such as the singular value decomposition (SVD) [45]. However, since they are conceptually identical, i.e., the outcome is a set of orthogonal vectors, we can expect performances similar to that of GS KAPE. GS KAPE method differs from most other methods in execution, since it iteratively operates over a set of vectors until the stopping criterion is met. This is in contrast to directly operating on the $MN \times N_p$ matrix of steering vectors. As we have demonstrated, the number of iterations in GS KAPE is much less than the number of space-time steering vectors N_P, and hence, is computationally efficient. In comparison, the often used SVD method is a two-step process that first reduces the input to a bidiagonal matrix. The second step is a QR decomposition, which is an iterative process over the $MN \times N_p$ bidiagonal matrix. Overall, the SVD is an involved process, requiring nontrivial calculations.

The SVD allows for a compact representation of the data. The number of significant singular values, i.e., singular values with magnitudes greater than some marginal value ϵ, indicates the minimal number of orthogonal subspaces required for a good approximation of the original matrix $\boldsymbol{V}$. Since GS KAPE also searches for orthogonal subspaces to represent the clutter ridge, we should expect the number of significant singular values to be the same as the number of GS KAPE output vectors. Figure 12 plots the magnitude of the singular values of $\boldsymbol{V}$ for the test system used to generate the SINR plots of Figure 10. We see that there are 37 significant singular values, which is equal to the number of orthogonal vectors obtained with GS KAPE.

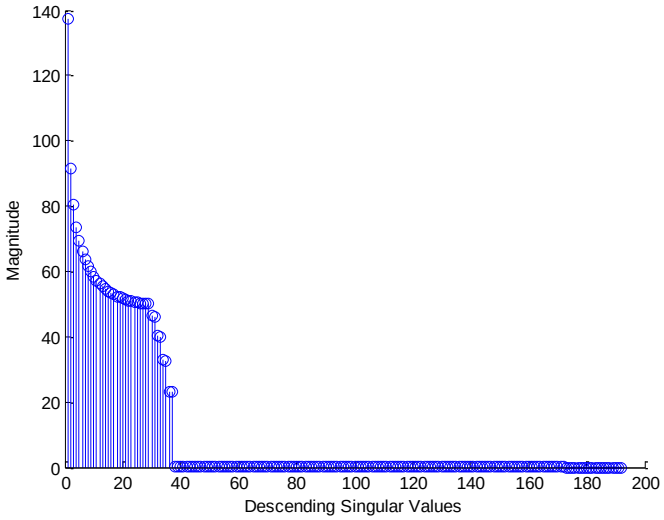


Figure 12. SVD singular values of a matrix of space-time steering vectors.

4.3. Power-Ridge Aligning Array Calibration

The SMI method indirectly calibrates complex gain mismatches, since angle-independent array errors persist across the range bins. That is, the received data vector for the l^{th} range bin is

$$\boldsymbol{x}_l = \sum_{k=1}^{N_p} \alpha_{l,k}[\boldsymbol{s}_k \odot (\boldsymbol{1}_N \otimes \boldsymbol{\epsilon})] = \sum_{k=1}^{N_p} \alpha_{l,k}\boldsymbol{v}_k,$$

where $\boldsymbol{\epsilon}$ is the actual array error, which is independent of the range bin l. The vectors $\boldsymbol{s}_1, \dots, \boldsymbol{s}_{N_p}$ represent the ideal space-time steering vectors. Let $\boldsymbol{V} = [\boldsymbol{v}_1, \dots, \boldsymbol{v}_{N_p}]$ and $\overline{\boldsymbol{P}} = \text{diag}\left\{\overline{|\alpha_1|^2}, \dots, \overline{|\alpha_{N_p}|^2}\right\}$, where $\overline{|\alpha_k|^2}$ denotes the power received from the k^{th} clutter patches, averaged over the available range bins. Then the SMI covariance estimate is

$$\boldsymbol{R}_{SMI} = \boldsymbol{V}\overline{\boldsymbol{P}}\boldsymbol{V},$$

which accounts for array errors. Note that this is a simplified model, as the actual array errors can jitter in gain and phase across the range bins. It nonetheless emphasizes the adaptability of SMI to array errors. In contrast, KAPE initially models the clutter ridge without assuming array errors:

$$\boldsymbol{R}_{KAPE} = \boldsymbol{S}\widehat{\boldsymbol{P}}\boldsymbol{S},$$

where $\boldsymbol{S} = [\boldsymbol{s}_1, \dots, \boldsymbol{s}_{N_p}]$ and $\widehat{\boldsymbol{P}}$ is a diagonal matrix of the power estimates. The angle-independent error term estimates $\hat{\boldsymbol{\epsilon}}$ are directly factored in upon applying a calibration algorithm, resulting in the corrected estimate

$$\boldsymbol{R}_{KAPE} = \widehat{\boldsymbol{V}}\widehat{\boldsymbol{P}}\widehat{\boldsymbol{V}},$$

where $\widehat{\boldsymbol{V}} = [\widehat{\boldsymbol{v}}_1, \ldots, \widehat{\boldsymbol{v}}_{N_p}] = [\boldsymbol{s}_1 \odot (\boldsymbol{1}_N \otimes \hat{\boldsymbol{\epsilon}}), \ldots, \boldsymbol{s}_{N_p} \odot (\boldsymbol{1}_N \otimes \hat{\boldsymbol{\epsilon}})]$. This underlies the importance of accurate calibration algorithms in the application of KAPE; array errors ultimately correspond to model mismatch.

For simulation, we generate a zero-mean, M-dimensional white Gaussian random variable $\boldsymbol{\chi}_1$ with standard deviation σ_1 that reflects the array gain error. The array phase error is represented with a second zero-mean, M-dimensional white Gaussian random variable $\boldsymbol{\chi}_2$ with standard deviation σ_2. The array error $\boldsymbol{\epsilon}$ is simulated by combing $\boldsymbol{\chi}_1$ and $\boldsymbol{\chi}_2$:

$$\boldsymbol{\epsilon} = (1 + \boldsymbol{\chi}_1) \odot \exp\{j\boldsymbol{\chi}_2\}.$$

Hence, the actual steering vectors are

$$\boldsymbol{v}_k = \boldsymbol{s}_k \odot (\boldsymbol{1}_N \otimes \boldsymbol{\epsilon}), \text{ for } k = 1, \ldots, N_p,$$

where $\boldsymbol{s}_k$ is the space-time steering vector for the k^{th} clutter patch and N_p is the total number of clutter patches describing the iso-ring.

To illustrate the effects of array errors, the dashed line of Figure 13 shows the SINR losses incurred by KAPE in the presence of uncalibrated array errors. Array gain errors with 3 dB standard deviation and array phase errors with 20 degrees standard deviation are simulated. We see that uncalibrated array errors lead to unacceptable performance with losses up to -38 dB around the 3rd Doppler bin. The dash-dot plot shows the improvement achieved when compensating using the adjacent-channel pair (ACP) method, as discussed in Section 2.6.1.

However, we still observe losses greater than -5 dB for select Doppler bins, which motivates the development of an improved array calibration technique. The result of applying KAPE with no array errors is also shown as a benchmark. Figure 14 plots the SINR losses incurred by SMI with uncalibrated array errors to demonstrate its adaptability. We see losses up -9.3 dB without array calibration, which is a vast improvement over KAPE in the same scenario. Upon applying the ACP method, the SINR suffers virtually no losses from the case with no array errors.

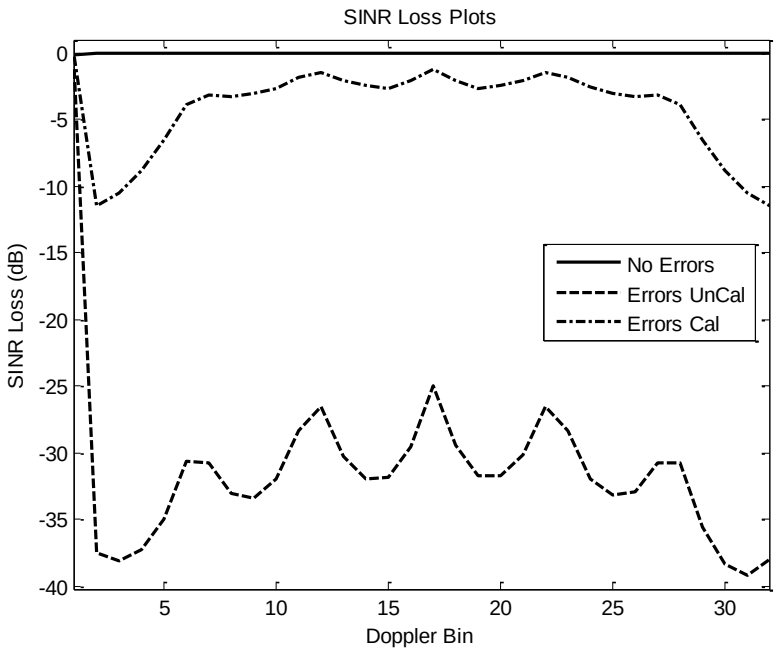


Figure 13. SINR loss plots demonstrating the losses incurred using the KAPE covariance matrix in the presence of array errors. Cases for compensated and uncompensated array errors are shown.

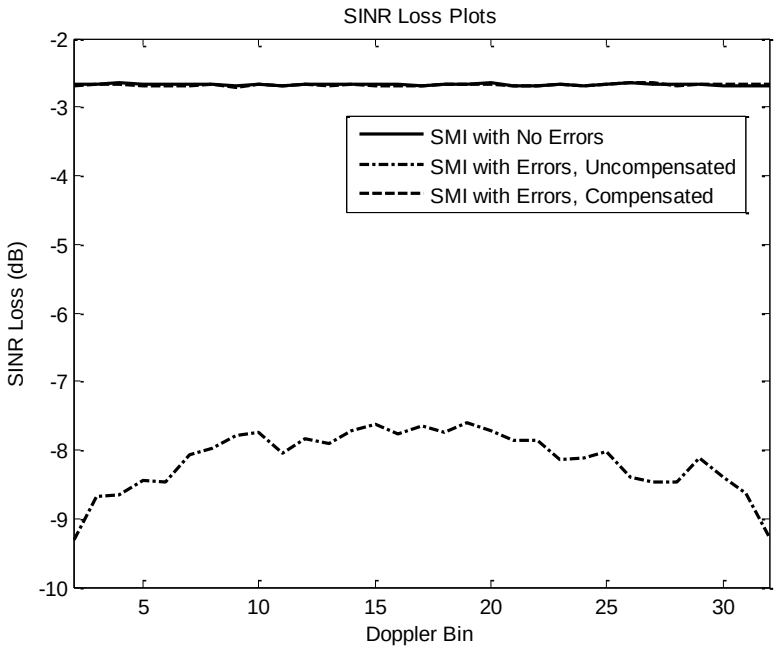


Figure 14. SINR loss plots demonstrating the losses incurred using the SMI covariance matrix in the presence of array errors. Cases for compensated and uncompensated array errors are shown.

We have developed an approach we call the power-ridge aligning (PRA) method, an iterative technique that searches for the best error vector $\hat{\boldsymbol{\epsilon}}$ to align the steering vectors to the clutter ridge of the data. We can rephrase the k^{th} term in the formulation of the clairvoyant clutter covariance matrix in terms of the temporal and spatial steering vectors

$$\boldsymbol{R} = \sum_{k=1}^{N_p} |\alpha_k|^2 \left(\boldsymbol{s}_{t/k} \otimes \left(\boldsymbol{s}_{s/k} \odot \boldsymbol{\epsilon}\right)\right)\left(\boldsymbol{s}_{t/k} \otimes \left(\boldsymbol{s}_{s/k} \odot \boldsymbol{\epsilon}\right)\right)^H,$$

where $\boldsymbol{s}_{t/k}$ and $\boldsymbol{s}_{s/k}$ are the $N \times 1$ temporal and $M \times 1$ spatial steering vectors, respectively, for the k^{th} clutter patch. We denote $|\alpha_k|^2$ as the power corresponding to the k^{th} patch. Note that the $M \times 1$ clairvoyant array error $\boldsymbol{\epsilon}$ only applies to the spatial steering vector. In

the case of no array errors, $\boldsymbol{\epsilon} = \mathbf{1}_M$. Letting $\boldsymbol{s}_o$ be the ideal space-time steering vector of the mainlobe, the calibrated space-time steering vector can be written as:

$$\boldsymbol{v}_o = \boldsymbol{s}_o \odot (\mathbf{1}_N \otimes \boldsymbol{\epsilon}).$$

It is easy to see that

$$\boldsymbol{s}_o^H \boldsymbol{R} \boldsymbol{s}_o \leq \boldsymbol{v}_o^H \boldsymbol{R} \boldsymbol{v}_o$$

due to the misalignment. That is, we obtain maximal power response when the space-time steering vector is aligned to the primary clutter subspace. This paves the way to define an error function to search for the estimate error vector $\hat{\boldsymbol{\epsilon}}$. Specifically, by forming the vector $\hat{\boldsymbol{v}}_o = \boldsymbol{s}_o \odot (\mathbf{1}_N \otimes \hat{\boldsymbol{\epsilon}})$, we define a simple error function

$$J(\hat{\boldsymbol{\epsilon}}) = \hat{\boldsymbol{v}}_o^H \hat{\boldsymbol{R}}^{-1} \hat{\boldsymbol{v}}_o \,,$$

where the covariance matrix is obtained through SMI. By using SMI, we ensure that the array errors are present in the estimate. The form of the error function resembles the minimum-variance distortionless beamfomer [41], a method to obtain super-resolved spectra.

In cases of interest, we do not have sufficient data to perform the matrix inversion required to calculate $\hat{\boldsymbol{R}}^{-1}$, which forces us to use dimension-reducing algorithms. In our simulations, we use the EFA algorithm outlined in Section 2.5.1. Denoting $\boldsymbol{T}$ as the dimension-reducing transformation matrix, we obtain the revised error function

$$J(\hat{\boldsymbol{\epsilon}}) = (\boldsymbol{T}\hat{\boldsymbol{v}}_o)^H (\boldsymbol{T}\boldsymbol{R}\boldsymbol{T})^{-1} (\boldsymbol{T}\hat{\boldsymbol{v}}_o) \,.$$

Many minimization algorithms could be tried. Here, we use the Conjugate Gradient Method [44], an iterative optimization technique that is able to converge to a solution within M iterations, where M is the dimension of the problem. Starting with an initial estimate $[1 \quad 0 \quad \cdots \quad 0]^T$, the gradient vector is estimated in the space defined by the error function. The step-size in the direction of the gradient is optimized and the estimate $\hat{\boldsymbol{\epsilon}}$ is updated. The conjugate gradient algorithm iteratively calculates subsequent conjugate gradients, which are gradients vectors orthogonal in the error space, along with the optimal descent step-sizes to further refine $\hat{\boldsymbol{\epsilon}}$. For an M-dimensional problem, there are exactly M conjugate vectors, and hence, the CG algorithm will converge within M steps.

We now compare the performance of PRA against other popular array calibration methods, namely the maximum eigenvector method and the adjacent channel pairing (ACP) method of Section 2.6.1. Figure 15 compares the mean-square errors (MSE) of the resulting array error estimates over $S = 1000$ samples on a uniform linear array. Specifically, we define the mean error for the k^{th} element to be

$$\frac{1}{S}\sum_{s=1}^{S}\left(\epsilon_{k,s} - \hat{\epsilon}_{k,s}\right)^2$$

over $M = 6$ elements. We define $\epsilon_{k,s}$ and $\hat{\epsilon}_{k,s}$ to be the actual and estimate errors of the k^{th} element and s^{th} sample, respectively. We simulated Gaussian array gain errors with a standard deviation of 3 dB and Gaussian array phase errors with a standard deviation of 20 degrees. Figure 15 shows the MSE for each element, where the first element is omitted because it is always 0, being the reference in the estimates. We see that PRA outperforms the ACP and the eigenvector methods by a significant margin. Unlike the ACP method, the

errors estimates do not increase as a function of the element number since it minimizes over the entire N dimensional space. The calculations of ACP method are performed in a pairwise fashion, sequentially across the linear array. Since the previous elemental estimates are used for reference, the errors cascade, explaining the performance loss as a function of the element number.

To better understand PRA's efficacy in clutter cancellation, we calculate the maximum eigenvalue of the product $\widehat{\boldsymbol{R}}^{-1}\boldsymbol{R}$, which gauges the mismatch between the calibrated KAPE covariance matrix and the clairvoyant covariance matrix. This measures the residue of the estimate, i.e., the power of the clutter that has not been suppressed. Again, we compare the PRA against the maximum eigenvector and the ACP methods over 1000 samples over varying levels of CNR. The results are shown in Figure 16. We see that PRA demonstrates a significant advantage, attaining 30 dB additional clutter cancellation compared to the adjacent channel method when the CNR is set to 50 dB.

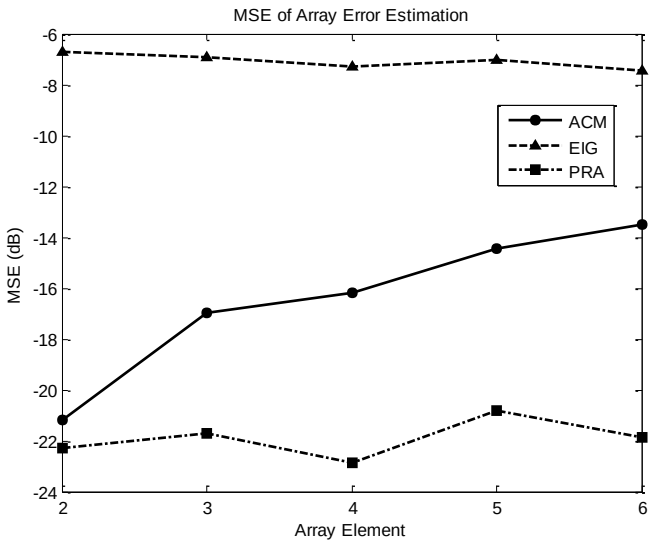


Figure 15. Mean-square error residue from various array calibration techniques.

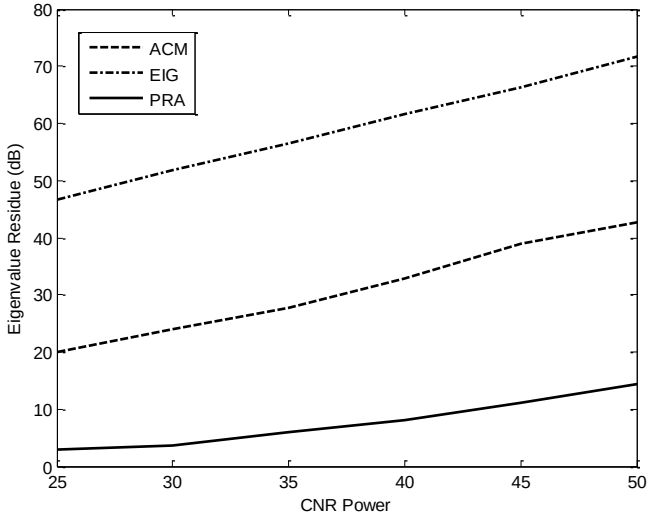


Figure 16. Plots showing the residue power due to mismatch in covariance matrix estimates.

4.4. E-KAPE Performance

We refer to the combination of PRA and GS-KAPE as enhanced KAPE (E-KAPE). This section compares the performances of E-KAPE against the standard KAPE and SMI methods. We first look at the homogeneous clutter case with the inclusion of array errors to verify that E-KAPE is on par with the standard SMI method. SMI performs well, as it is designed to perform in homogeneous clutter and naturally compensates for array errors through averaging. We then introduce heterogeneities to the data, namely point clutter as discussed in Section 2.3.2. The residues of insufficiently nulled clutter discretes (CD) can trigger false alarms, while targets in secondary data (TSD) form nulls within the target space, leading to missed detections.

To illustrate the effects of point clutter in the data, example range-Doppler maps are generated for various cases of interest. Range-Doppler maps are power responses of the received data $\boldsymbol{x}_l,\ l = 1, \dots, L$ corresponding to Doppler bins $n = 1, \dots, N$ at fixed azimuth

and elevation angles. Specifically, the range-Doppler map value of the l^{th} range bin and n^{th} Doppler bin is calculated as

$$p_{l,n} = |\boldsymbol{s}(f_n, \theta_o)^H \boldsymbol{x}_l|^2,$$

where $\boldsymbol{s}(f_n, \theta_o)$ is the space time steering vector corresponding to the n^{th} Doppler bin at a fixed azimuth angle θ_o. The elevation is set to 0^o for the examples. Figure 17 demonstrates an example range-Doppler map in homogeneous clutter, illustrating the ground clutter centered about the 0^{th} Doppler bin; this corresponds to the clutter that STAP filters are designed to suppress. In Figure 18, a clutter discrete is injected into range bin $l = 27$. Since clutter discretes have zero velocities, it is centered about DC. However, due to their high power-to-noise ratio, we see severe spectral leakage across the Doppler bins. If unaccounted for, the leaked power can be wrongfully classified as a target. Figure 19 demonstrates the effect of TSDs in an example range-Doppler map. In this case, a TSD with radial velocity $v = 40\ \mathrm{m/s}$ at azimuth $\theta = -5.25^o$ is injected into range bin $l = 36$. If the SMI covariance estimate is used to calculate the STAP weights, a null will form at the angle-Doppler location of the prominent TSD, which will consequently filter the corresponding targets of interest. Figure 20 plots the resulting SMI covariance matrix, where the exemplified TSD is seen to be centered around 0.71 normalized Hertz.

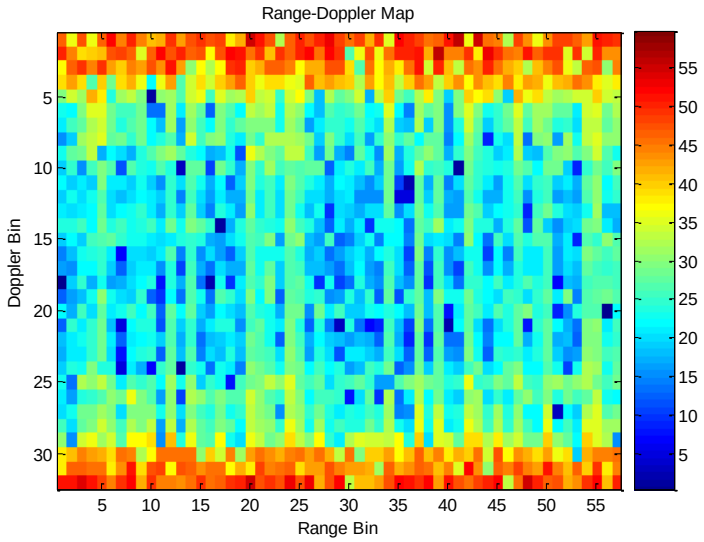


Figure 17. An example of a range-Doppler map in homogeneous clutter.

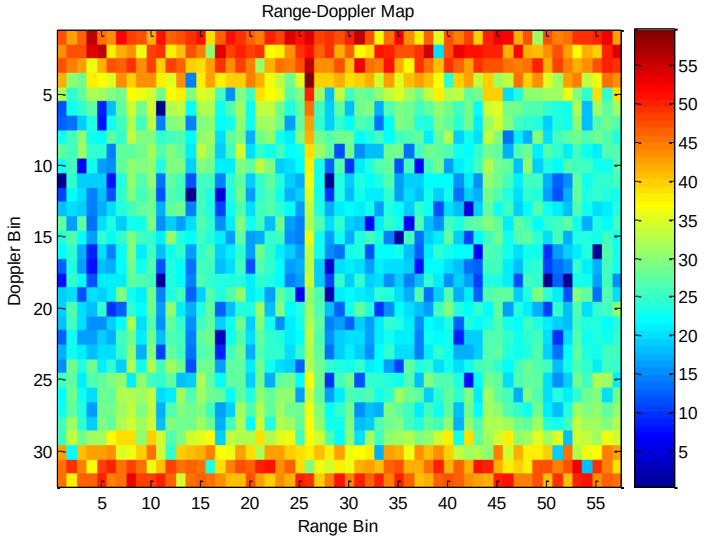


Figure 18. An example of a range-Doppler map in the presence of clutter discretes (CD).

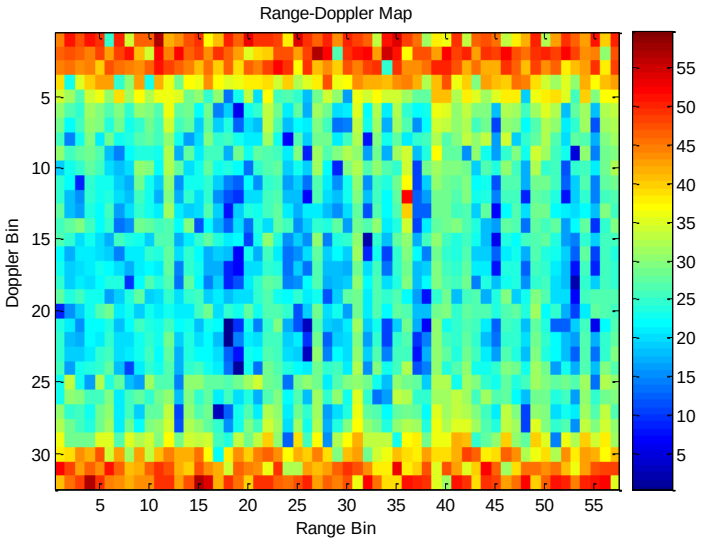


Figure 19. An example of a range-Doppler map in the presence of targets in secondary data (TSD).

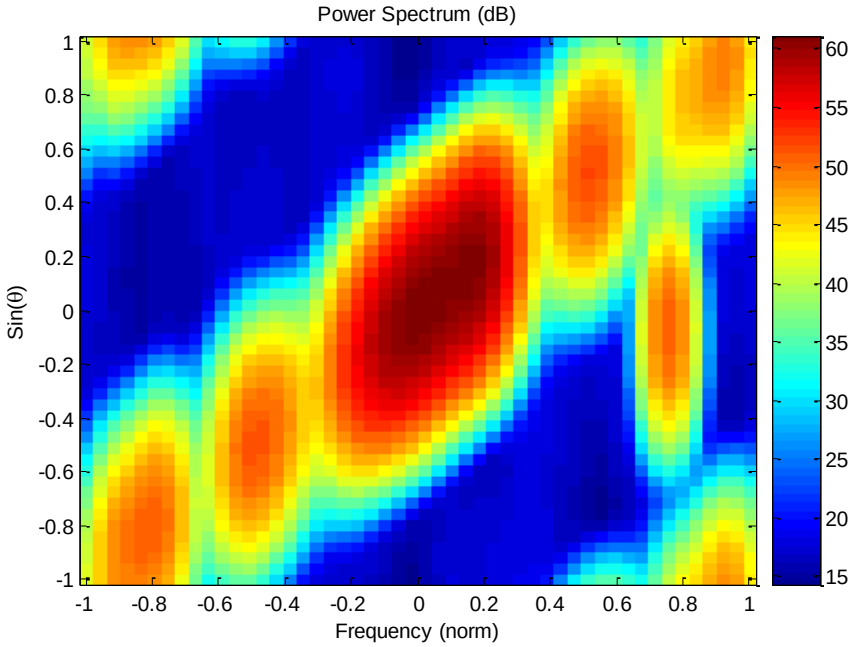


Figure 20. An example of an SMI covariance estimate contaminated with TSD.

4.4.1. Homogeneous Clutter

The simulated, homogeneous clutter environment is characterized by clutter with 25 dB single-channel, single-pulse (SCSP) clutter-to-noise ratio (CNR). The radar system is afflicted by Gaussian array gain errors with a standard deviation of 3 dB and Gaussian array phase errors with a standard deviation of 20 degrees. The receiver consists of $M = 6$ elements and transmits $N = 32$ pulses per CPI. We have $L = 36$ range bins available for secondary data. The SMI method uses EFA for dimension reduction with 3 sub-pulses; the E-KAPE and KAPE methods do not require this step to perform the matrix inversion. SMI and KAPE perform array calibration via the adjacent channel pairing (ACP) method, while E-KAPE employs our power-ridge aligning (PRA) method. All results are averaged over 1000 trials.

SINR loss curves in Figure 21 demonstrate that E-KAPE outperforms SMI by about 3 dB for most Doppler bins, the amount SMI loses due to averaging. KAPE is hit with large losses due to inadequate calibration via the ACP method. In particular, we observe SINR losses over -11 dB for Doppler bins around the "direct current" (DC) component. For the same set of Doppler bins, we attain a maximum loss of -1 dB with E-KAPE through the use of the PRA method, which is a huge improvement. This is expected, as our analysis in Figure 16 of Section 4.3 shows that the PRA method appears to suppress 18 dB of additional eigenvalue residue power in a 25 dB CNR scene. Note that in the absence of array errors, we expect E-KAPE and KAPE to perform similarly, both outperforming SMI. This was demonstrated in Figure 10 of Section 4.2, where we compared SINR loss plots to demonstrate the efficacy of GS KAPE.

The ROC curves of Figure 22 reflect what we observed from the SINR loss plots. In contrast to Section 3.2, which demonstrated excellent results with KAPE in homogeneous clutter, we see that KAPE underperforms SMI because of the insufficiently calibrated array errors. E-KAPE employs the PRA method and does not suffer the same losses, as effectively shown in Figure 22. For a different perspective on the problem, the exceedance plots are shown in Figure 23. We can see the advantage with both KAPE methods: a lower threshold is needed to obtain specific probability of false alarms, which generally improves the probability of detection. However, the differences in the required thresholds are small, so this does not fully explain the gaps in the ROC curves. Figure 24 plots the probability of detections obtained with the SMI, KAPE, and E-KAPE methods as a function of the AMF threshold used. E-KAPE maintains higher detection probabilities for the thresholds compared to the other methods. As an example, to maintain $P_D = 0.9$, E-KAPE requires a threshold of approximately $T = 6.01$. In contrast, SMI and KAPE needs to lower the threshold to $T = 3.3$ and $T = 1.33$, respectively. Lowering the threshold increases the false alarm rate, which leads to poor ROC performance.

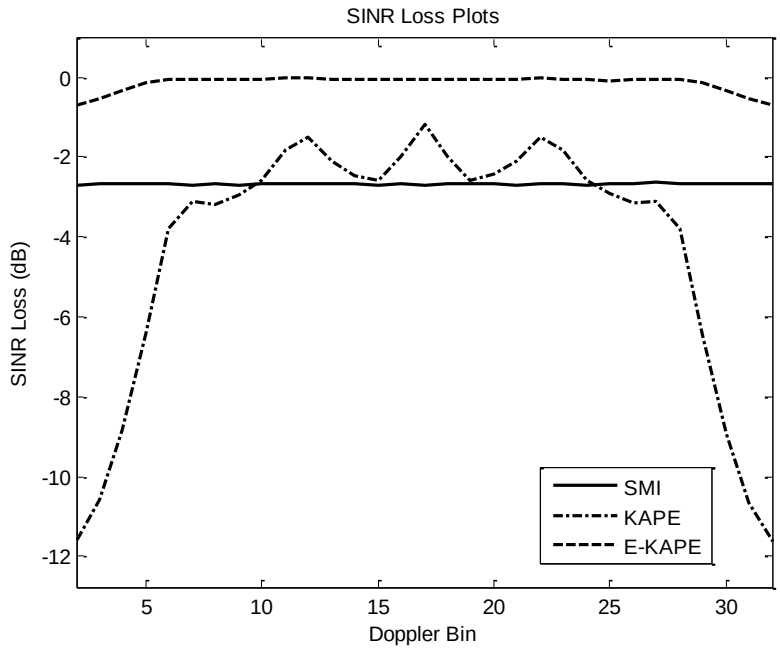


Figure 21. SINR plots comparing various STAP methods in homogeneous clutter.

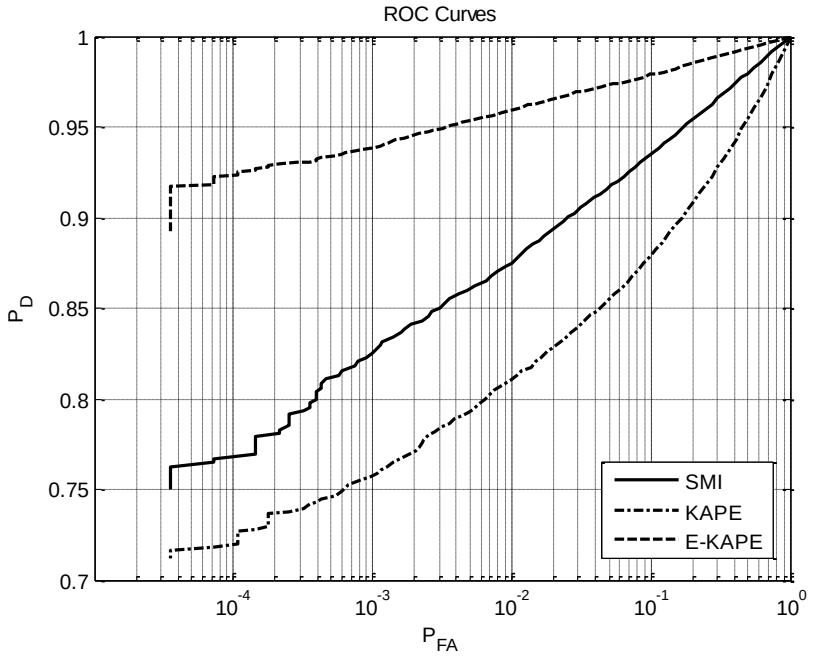


Figure 22. ROC curves comparing various STAP methods in homogeneous clutter.

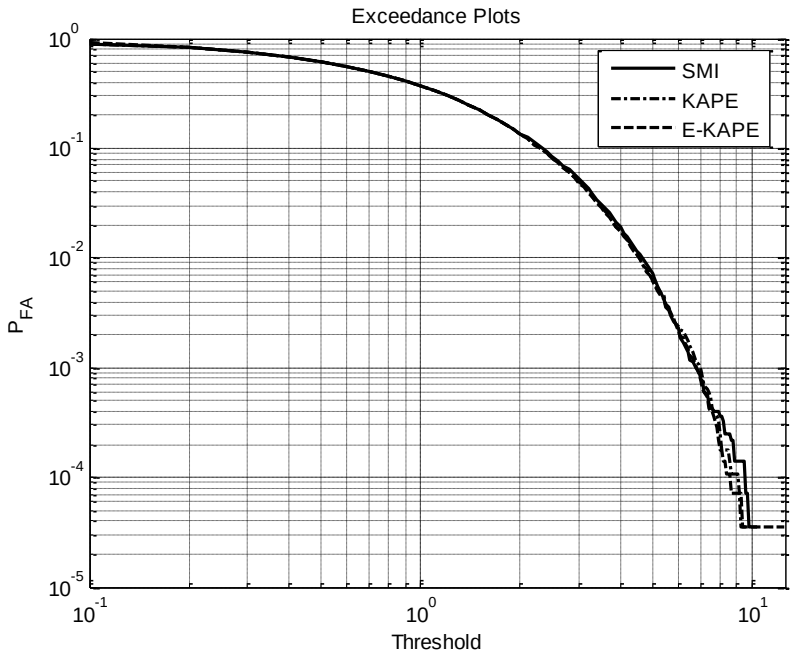


Figure 23. Exceedance plots comparing various STAP methods in homogeneous clutter.

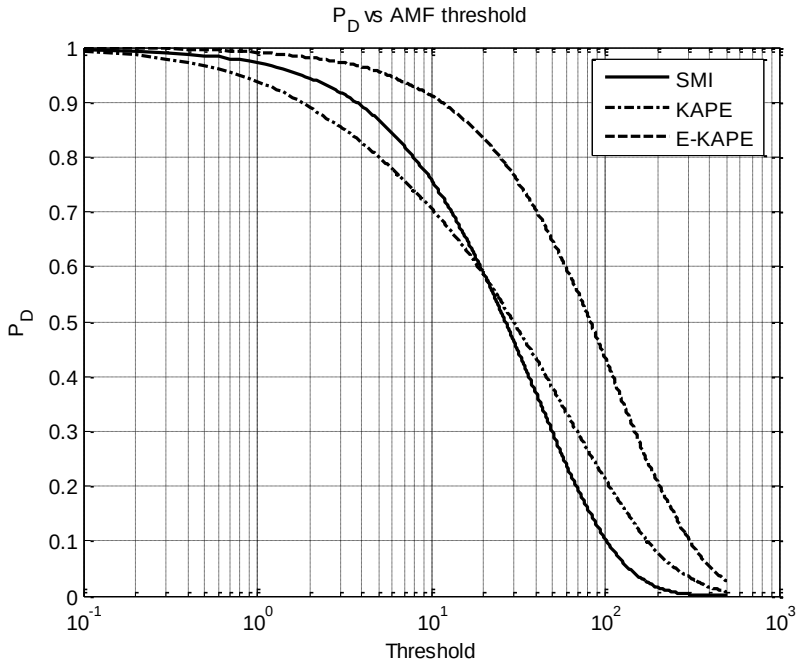


Figure 24. Probability of detection versus the AMF threshold for the SMI, KAPE, and E-KAPE methods.

4.4.2. Heterogeneous Clutter

For the heterogeneous case, we seed CDs and TSDs onto the homogeneous background with 25 dB SCSP CNR. We inject CDs of 45 dB SCSP power, uniformly distributed between 4 to 10 degrees from the mainlobe. The number of CDs in each trial is chosen randomly based on a Poisson distribution with mean $\lambda = 1$. We simultaneously inject TSDs of 30 dB SCSP power, traveling at 30 m/s and uniformly distributed over -10 to -4 degrees from the mainlobe. In addition to the point clutter, we include 3 dB gain and 20 degrees phase Gaussian array errors in the system. The same radar system is employed, i.e., $M = 6$, $N = 32$, and $L = 36$, with results averaged over 1000 trials.

Figure 25 plots the SINR losses of the STAP systems incurred in heterogeneous clutter. As expected, compared to the homogeneous case, the introduction of the extra complexities reduces the performance of all three STAP methods. In particular, we see an SINR loss of approximately -0.2 dB for E-KAPE across the Doppler bins. Overall, E-KAPE comes out on top compared to KAPE and the SMI methods. E-KAPE outperforms SMI by about 3 dB for the Doppler bins away from the 7^{th} bin. The dip at the 7^{th} Doppler bin is the precise frequency that corresponds to the seeded TSDs. Recall that unsuppressed TSDs contaminate the covariance estimate, as discussed in Section 4.4 and illustrated in Figure 20. The KAPE methods avert this loss by only parameterizing along the clutter ridge, ignoring the TSD residues off the clutter ridge. However, KAPE is plagued with large losses due to inadequate array calibration, resulting in losses of almost to -12 dB for Doppler bins around the DC component.

Figure 26 shows the ROC plots that reflect the relative performances observed from the SINR analysis. By employing the PRA method, E-KAPE avoids the additional losses KAPE suffers from insufficiently calibrated array errors. A huge dip in ROC performance is seen for the SMI case due to the presence of CDs. CDs are essentially strong outliers, requiring large AMF thresholds to suppress them. The increase of the thresholds required to suppress the discretes significantly diminishes the probability of detection.

The exceedance plots shown in Figure 27 better demonstrate the effects of clutter discretes. The SMI method does not sufficiently suppress the injected CDs, which is evidenced by the long tail of the exceedance curve. In contrast, the E-KAPE method does an excellent job since only the CUT is required to generate the covariance matrix estimate. It is unaffected by adjacent range bins that may not contain CD spikes. Figure 28 offers a zoomed exceedance plot, accentuating the AMF thresholds needed to maintain the desired probability of false alarms. In particular, to maintain false alarm rate of 5×10^{-4}, E-KAPE requires a threshold of 8.12, whereas SMI requires a threshold of 13.2. The higher threshold needed for the SMI method implies a decrease in the probability of detection, as demarcated in Figure 29. For E-KAPE, an AMF threshold of 8.12 results in $P_D = 0.87$. The threshold of 13.2 required for the SMI method results in $P_D = 0.68$, which is a huge loss in performance.

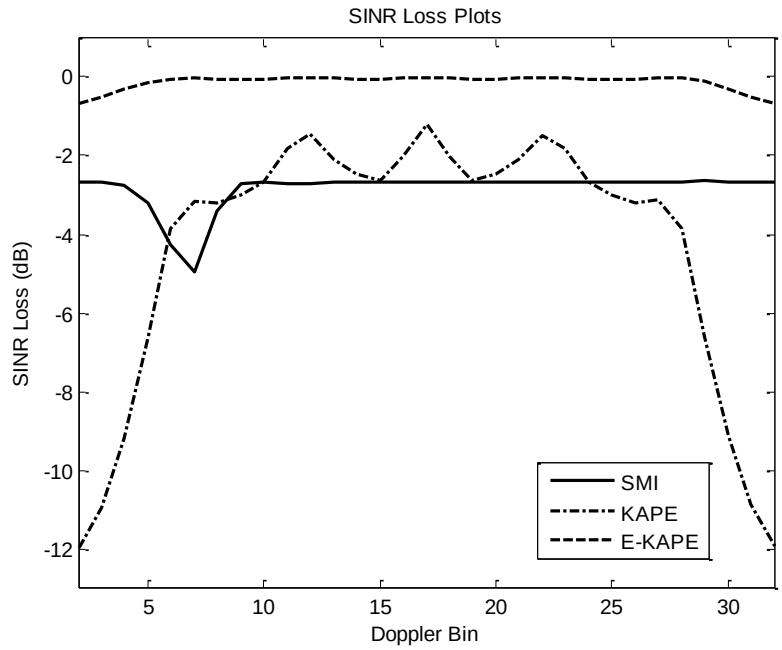


Figure 25. SINR plots comparing various STAP methods in heterogeneous clutter.

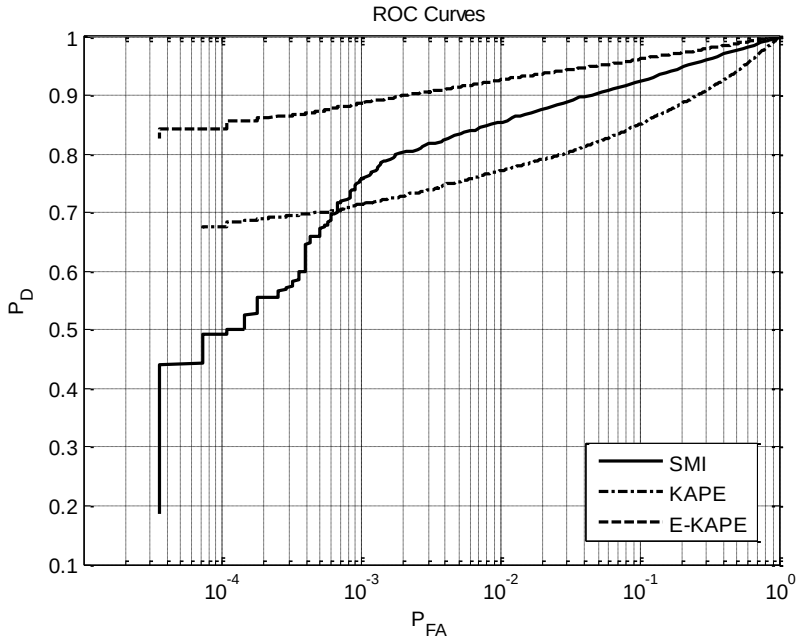


Figure 26. ROC curves comparing various STAP methods in heterogeneous clutter.

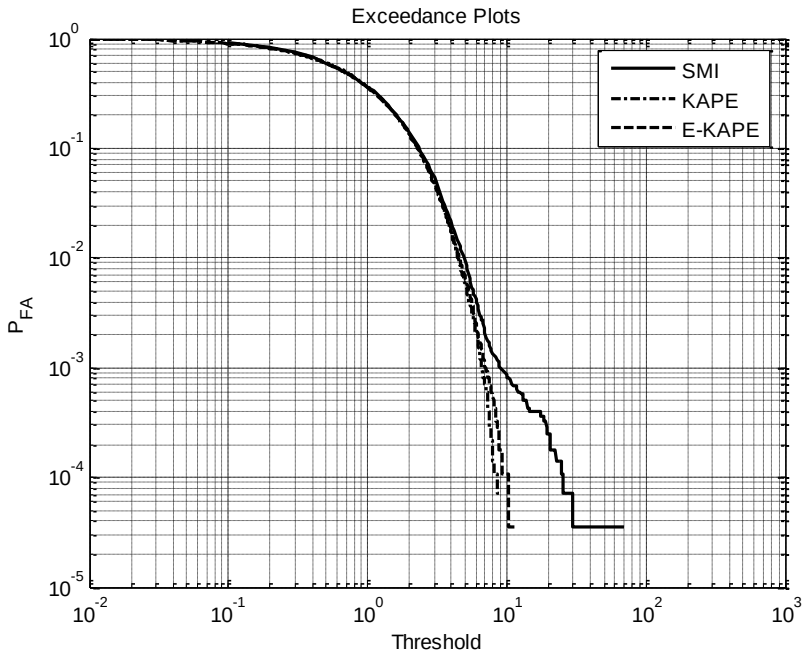


Figure 27. Exceedance plots comparing various STAP methods in heterogeneous clutter.

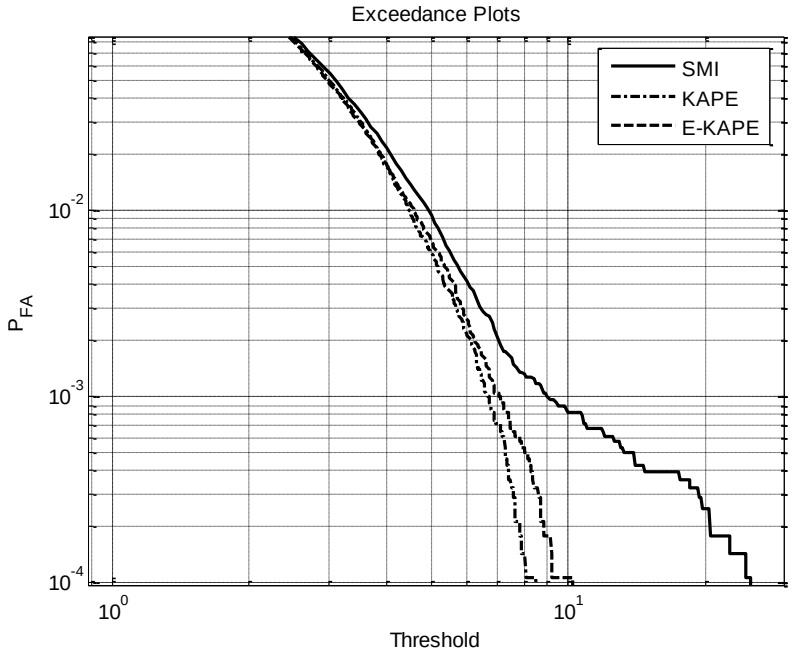


Figure 28. Exceedance plots comparing various STAP methods in heterogeneous clutter.

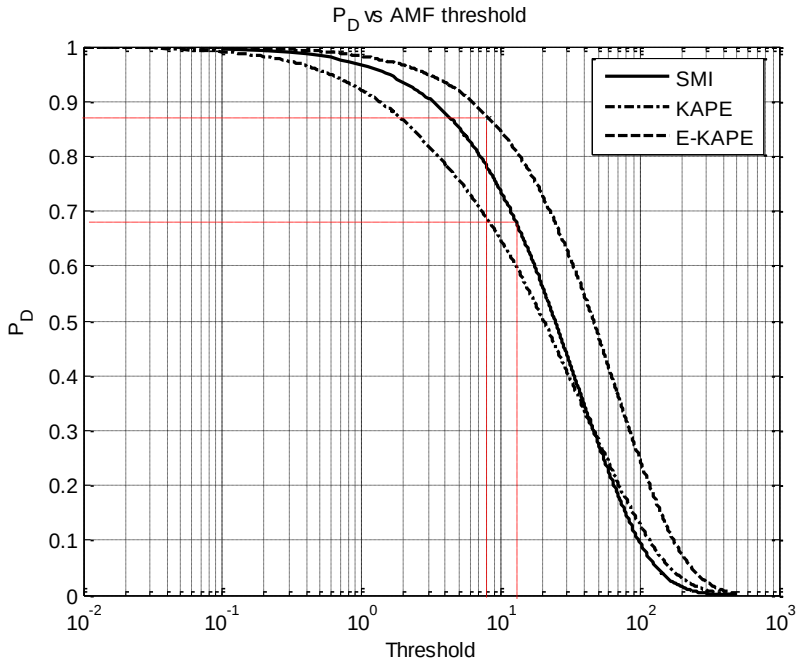


Figure 29. Plot of attainable probability of detection in the presence of point clutter and array errors for SMI, KAPE and E-KAPE methods.

4.5. Computational Complexity of E-KAPE and KAPE

This section analyzes the differences in the computational requirements of KAPE and E-KAPE in the simulations discussed in Section 4.4. The dimension of the data is $MN = 192$, implying that the clutter covariance matrices $\boldsymbol{R}$ are of size 192×192. The two modules in which the KAPE methods differ include the array calibration step, the parameterization of the covariance matrix, and the calculation of its subsequent inverse.

We first compare calibration methods. The main computational requirements for the ACM are in the calculation of the adjacent covariance matrices. This amounts to $N - 1$ matrix multiplications between the $2 \times M$ and $M \times 2$ channel matrices. The PRA employs the conjugate gradient (CG) algorithm, which is an iterative technique, making it hard to

ascertain the precise computational requirements. The CG is performed over the reduced space, the result of applying EFA with $K = 3$ subpulses. Hence, the complexity for every iteration is $O(KN)$. The number of iterations depends on the stopping criteria and the conditions of the CUT. In general, PRA is more computationally complex than the ACM.

Now, we compare the complexities of the covariance matrix parameterizations. E-KAPE starts off by calling on the modified Gram Schmidt (GS) algorithm to obtain a reduced, orthonormal set of space-time steering vectors. The exact computational requirements for the GS procedure depend on the system parameters, namely the input space-time steering vectors. It is difficult to calculate this ahead of time since it is an iterative algorithm, looping until the stopping criteria is met. In general, it is not an involved step, only requiring vector multiplications. For our simulations, the modified GS outputs 37 space-time steering vectors, which equates to 38 iterations; the last iteration does not produce an output. Hence, for our simulated environments, E-KAPE only requires 37 power estimates to form the estimate covariance matrix, versus the 656 estimates required for the oversampled KAPE. Although the power estimates are required for each range sample, the GS algorithm only needs to be applied once. The largest computational difference is in the calculation of the inverse covariance matrix. The standard KAPE algorithm requires the inversion of a 192×192 matrix, requiring a computational cost of $O(M^3N^3)$. This itself can be prohibitive for real systems, and hence, KAPE is typically implemented in a reduced-dimension space. In contrast, E-KAPE employs a parameterized form of the inverse covariance matrix, where the crux of the computation stems from the outer products between the 37 orthonormal vectors. This difference enables E-KAPE to be used in practice.

Chapter 5

CONCLUSION

We conclude by reviewing the main contributions of this paper, which includes the innovation of the cell-based clutter model (CCM), Gram-Schmidt (GS) KAPE, and the power ridge aligning (PRA) array calibration algorithm. This is followed by discussing directions for further research.

5.1. Contributions

STAP techniques operate in the space and time dimensions to effectively suppress ground clutter and extract moving targets. Many well-known techniques, such as the SMI, operate under the assumption that the clutter is homogeneous. This dissertation introduced various forms of heterogeneity, namely heterogeneous texture and point clutter, that invalidates this assumption. In the presence of heterogeneous texture clutter, we have demonstrated that a 0.5 loss in detection probability can be expected for a system that maintains a false alarm rate of 1×10^{-4}. This necessitates the development of specialized algorithms. To further validate our case, we have demonstrated a 0.3 loss in detection probability in the presence of point clutter for standard STAP.

The first step in developing improved algorithms is to obtain faithful examples of the clutter we aim to suppress. This underlines the importance of building an accurate clutter model in

which to simulate our data. We have developed the cell-based clutter model (CCM), which has the ability to generate clutter scenes with heterogeneous texture as well as point clutter. Remarkably, the CCM is able to generate a gamut of heterogeneous texture scenes with two control parameters. This allowed us to test STAP algorithms on scenes ranging from near homogeneous to highly volatile, spiky scenes. With this in hand, we have proved the effectiveness of KAPE in heterogeneous texture clutter; the efficacy of KAPE did not diminish since it is allowed to adapt solely on the cell-under-test. Even in homogeneous clutter, we have shown that we actually gain 0.1 improvement in detection probability over the SMI algorithm for a system that maintains a false alarm rate of 1×10^{-4}.

Two problems hinder KAPE's implementation in real systems: its computational complexity and its susceptibility to array errors. The crux of the computational burden lies in the direct inversion of $MN \times MN$ matrices, one for each range bin, which is a requirement in the calculation of the STAP weights. For this purpose, we have developed a Gram-Schmidt (GS) KAPE that circumvents the need for direct inversions, opting instead for a computationally simpler parametric form. Furthermore, GS KAPE reduces the overall number of steering vectors describing the clutter ridge, hence reducing the number of required power estimates. Through the CCM, we have demonstrated that these modifications do not result in significant losses.

We demonstrated the crippling effects of array errors when applying KAPE; such errors correspond to a model mismatch. We showed SINR losses of -35 dB with uncalibrated arrays and errors up to -10 dB with current calibration algorithms. To combat this shortcoming, we developed the power ridge aligning (PRA) calibration technique, an

iterative gradient descent technique, which is a new approach to the problem. Using PRA, we were able to maintain a -1 dB SINR loss in the presence of array errors.

5.2. Suggestions for Future Work

One problem we have not fully addressed is the compensating of intrinsic clutter motion (ICM) in the KAPE covariance estimate. The time correlations in the data cause the clutter ridge to spread in Doppler frequency. Since the steering vectors of KAPE lie strictly on the ridge, the unsuppressed off-diagonal clutter can lead to target masking. In standard KAPE implementation, ICM is compensated for by directly applying a covariance matrix taper (CMT) that matches the time correlation. The same approach cannot be used with the GS KAPE method, since the inverse covariance matrix is calculated parametrically.

We simulated ICM using the Gaussian temporal autocorrelation CMT described in Section 2.6.4 with a decaying parameter α. Figure 30 demonstrates the losses incurred for $\alpha = 2\ \mathrm{m}^2/\mathrm{s}^2$. To properly compensate for clutter spreading using GS KAPE, we require that the steering vectors span Doppler-angle bins slightly away from the clutter ridge. Our current approach initially takes MN steering vectors equally spaced on the clutter ridge to form an $MN \times MN$ matrix $\boldsymbol{S}$. An estimate of the spreading is formed using a Gaussian ICM CMT $\boldsymbol{C}(\beta)$ with parameter β that is distinct from α of the actual ICM. Before decomposing the steering vectors via the ordered Gram Schmidt algorithm, we directly apply the estimate spreading:

$$\boldsymbol{S} \mapsto \boldsymbol{S} \odot \boldsymbol{C}(\beta).$$

This gives us extra orthogonal components of the clutter ridge, which are needed for parametric inversion. Figure 30 demonstrates preliminary results using the method, where the spreading variable β was obtained empirically. Future studies could build on this idea to further lessen the SINR losses.

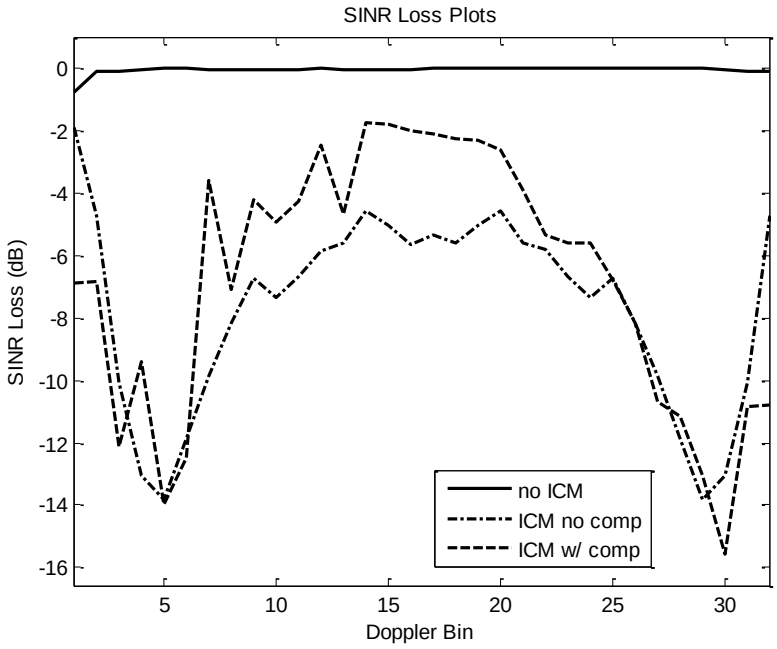


Figure 30. Losses incurred using KAPE in the presence of ICM.

References

[1] Bang, J., Melvin, W.L., Lanterman, A.D., "STAP Performance in K-Distributed Clutter," *2010 IEEE International Radar Conference*, pp. 1089-1093, May 2010.

[2] Bang, J., Melvin, W.L., Lanterman, A.D., "KAPE in K-Distributed Clutter," To be submitted to IET Signal Processing.

[3] Bang, J., Melvin, W.L., Lanterman, A.D., "Model-Based Clutter Cancellation Based on Enhanced Knowledge-Aided Parametric Covariance Estimation," To be submitted to IEEE Transactions on Aerospace and Electronic Systems.

[4] Richards, M.A., Fundamentals of Radar Signal Processing. McGraw-Hill, 2005.

[5] Ward, J., "Space-time adaptive processing for airborne radar," Lincoln Laboratory, Technical Report ESC-TR-94-109, Dec. 1994.

[6] Proakis, J.G. , Digital Signal Processing: Principles, Algorithms and Applications. Prentice Hall, 1995.

[7] Reed, I.S., Mallett, J.D., Brennan, L.E., "Rapid Convergence Rate in Adaptive Arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-10, no. 6, pp. 853-863, Nov. 1974.

[8] Ward, K.D., Baker, C.J., Watts, S. , "Maritime surveillance radar part 1: radar scattering from the ocean surface," In *Proceedings of the IEEE*, vol. 137, no. 2, pp. 51-62, April 1990.

[9] Ward, K.D., "Compound representation of high resolution sea clutter," *Electronic Letters*, vol. 17, issue 16, pp. 561-563, Aug. 1981.

[10] Dong, Y., "Models of land clutter vs. grazing angle, spatial distribution and temporal distribution : L-band VV polarization perspective," Electronics and Surveillance Research Laboratory, Report DSTO-RR-0273, March 2004.

[11] Nitzberg, R., Radar Signal Processing and Adaptive Systems. Artech House Publishers, 1999.

[12] Barton, D.K., "Land clutter models for radar design and analysis," In *Proceedings of the IEEE*, vol. 73, issue 2, pp. 198-204, Feb. 1985.

[13] Papoulis A., Pillai, S.U., Probability, Random Variables and Stochastic Processes, McGraw-Hill Europe, 2002.

[14] Billingsley, J. B., Low-Angle Radar Land Clutter: Measurements and Empirical Models, William Andrew Publishing, 2002.

[15] Sangston, K.J., Gerlach, K.R., "Coherent detection of radar targets in a non-Gaussian background," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, issue 2, Apr. 1994, pp. 330 – 340.

[16] Trunk, G., George, S., "Detection of targets in non-Gaussian sea clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-6, issue 5, 1970, pp. 620-628.

[17] Skolnik, M. I., Introduction to Radar Systems, McGraw-Hill, 2002.

[18] Watts, S., "Radar detection prediction in K-distributed sea clutter and thermal noise," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-23, issue 1, pp. 40-45, 1987.

[19] Lewinski, D., "Nonstationary probabilistic target and clutter scattering models," *IEEE Transactions on Antennas and Propagation*, vol. 31, issue 3, pp. 490-498, May 1983.

[20] Conte, E., Lops, M., Ricci, G., "Fitting the exogenous model to measured data", *IEEE Transactions on Instrumentation and Measurement*," vol. 43, issue 5, pp. 758-763, Oct 1994.

[21] Melvin, W.L., "Space-time adaptive radar performance in heterogeneous clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, issue 2, Apr. 2000, pp. 621-633.

[22] Melvin, W.L., Guerci, J.R., "Adaptive detection in dense target environments," In *Proceedings of the 2001 IEEE Radar Conference*, pp. 187-192, 2001.

[23] Swerling, P., "Probability of detection for fluctuating targets," *IRE Transactions on Information Theory*, vol. 6, issue 2, pp. 269-308, April 1960.

[24] Kelly, E.J., "An Adaptive Detection Algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-22, issue 2, pp. 115-127, March 1986.

[25] Scharf, L.L., McWhorter, L.T., "Adaptive matched subspace detectors and adaptive coherence estimators," In *Proceedings of the 30th Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1114-1117, Nov. 1996.

[26] Robey, F.C., Fuhrmann, D.R., Kelly, E.J., Nitzberg, R., "A CFAR adaptive matched filter detector," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, issue 1, pp. 208-216, Jan. 1992.

[27] Gabriel, W., "Using spectral estimation techniques in adaptive processing antenna systems," *IEEE Transactions on Antennas and Propagation*, vol. 34, issue 3, pp. 291-300, March 1986.

[28] Tufts, D.W., Kirsteins, I., Kumaresan, R., "Data-adaptive detection of a weak signal," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-19, issue 2, pp. 313-316, March 1983.

[29] Kirsteins, I.P., Tufts, D.W., "Adaptive detection using low rank approximation to a data matrix," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, issue 1, pp. 55-67, Jan. 1994.

[30] DiPietro, R., "Extended factored space-time processing for airborne systems," In *Proceedings of the 26th Asilomar Conference on Signals, Systems, and Computing*, pp. 425-430, Oct 1992.

[31] Wang, H., Cai., L., "On adaptive spatial-temporal processing for airborne surveillance radar systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, no. 3, pp. 660-669, July 1994.

[32] Carlson, B.D., "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 24, issue 4, pp. 397-401, July 1988.

[33] Hiemstra, J.D., "Colored diagonal loading," In *Proceedings of the 2002 IEEE Radar Conference*, pp. 386-390.

[34] Guerci, J.R., "Theory and application of covariance matrix tapers for robust adaptive beamforming," *IEEE Transactions on Signal Processing*, vol. 47, issue 4, pp. 977-985, April 1999.

[35] Melvin, W.L., Wicks, M.C., Brown, R.D., "Assessment of multichannel airborne radar measurements for analysis and design of space-time processing architectures and algorithms", In *Proceedings of the IEEE National Radar Conference*, pp. 130-135, May 1996.

[36] Rabideau, D.J., Steinhardt, A.O., "Improved adaptive clutter cancellation through data-adaptive training", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, issue 3, pp. 879-891, July 1999.

[37] Kogon, S.M., Zatman, M.A., "STAP adaptive weight training using phase and power selection criteria", In *Proceedings of the 35th Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 98-102, 2001.

[38] Melvin, W.L., Guerci, J.R., "Knowledge-aided signal processing: a new paradigm for radar and other advanced sensors," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, issue 3, pp. 983-996, July 2006.

[39] Melvin, W.L., Showman, G.A., "An approach to knowledge-aided covariance estimation", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, issue 3, pp. 1021-1042, July 2006.

[40] Guerci, J.R., Baranoski, E.J., "Knowledge-aided adaptive radar at DARPA: an overview," *IEEE Signal Processing Magazine*, vol. 23, issue 1, pp. 41-50, January 2006.

[41] Johnson, D.H., Dudgeon, D.E., Array Signal Processing: Concepts and Techniques. Prentice Hall, 1993.

[42] Melvin, W.L., Showman, G.A., Guerci, J.R., "A knowledge-aided GMTI detection architecture," In *Proceedings of the 2004 IEEE Radar Conference*, pp. 301-306.

[43] Skofronick-Jackson, G.M., Gasiewski, A.J., "A nonlinear multispectral statistical CLEAN-based precipitation parameter-retrieval algorithm ", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 1, pp.226-237, Jan. 2000.

[44] Press, W.H., Numerical Recipes in C: The Art of Scientific Computing. Cambridge University Press, 1992.

[45] Golub, G.H., Van Loan, C.F., Matrix Computations. The John Hopkins University Press, 1996.