AFRL-AFOSR-VA-TR-2022-0499

Computational and Statistical Foundations of Radar Signal Processing in the Low False Alarm Regime

Douglas Cochran
ARIZONA STATE UNIVERSITY
660 S MILL AVE STE 312
TEMPE, AZ, 85281
USA

08/10/2022
Final Technical Report

Air Force Research Laboratory
Air Force Office of Scientific Research
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED	
20220810	Final	START DATE 20180501	END DATE 20220430

4. TITLE AND SUBTITLE
Computational and Statistical Foundations of Radar Signal Processing in the Low False Alarm Regime

5a. CONTRACT NUMBER	5b. GRANT NUMBER	5c. PROGRAM ELEMENT NUMBER
	FA9550-18-1-0190	61102F
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**

6. AUTHOR(S)
Douglas Cochran

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
ARIZONA STATE UNIVERSITY
660 S MILL AVE STE 312
TEMPE, AZ 85281
USA

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)	11. SPONSOR/MONITOR'S REPORT NUMBER(S)
Air Force Office of Scientific Research 875 N. Randolph St. Room 3112 Arlington, VA 22203	AFRL/AFOSR RTB1	AFRL-AFOSR-VA-TR-2022-0499

12. DISTRIBUTION/AVAILABILITY STATEMENT
A Distribution Unlimited: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This project has addressed two foundational problems in statistical signal processing in the context of multi-channel radar: (1) Signal detection in the low false alarm regime of operation and (2) Estimation and tracking of signal subspaces. Prior to this project, work of the PI and others established that optimal multi-channel detection statistics take the form of functions of the eigenspectrum of a Gram matrix formed from data collected at the multiple receivers. This project focused on the most important special case for radar applications, where the signal has rank one and the key detection statistic is the largest eigenvalue of the Gram matrix. Setting detection thresholds that yield desired probabilities of false alarm requires explicit calculation of values of the conditional probability distribution of this statistic under the null hypothesis that the channels are statistically independent and contain only zero-mean white gaussian noise, in which case the Gram matrix is has a central Wishart distribution. Prior to the outset of this project, it had been observed that the probability distributions of the largest eigenvalue of such a matrix , although known analytically, were not amenable to direct numerical evaluation -- particularly for values of the statistic corresponding to low false-alarm probabilities necessary in practical radar signal detection scenarios and for realistic signal vector dimensionality and numbers of receiver channels. This project developed a suitable transformation of the distribution of the largest eigenvalue of a central Wishart distribution that circumvents the key obstructions to direct numerical calculations when the data dimensionality is large and the number of channels is in a realistic range.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U	UU	22

19a. NAME OF RESPONSIBLE PERSON	19b. PHONE NUMBER *(Include area code)*
ARJE NACHMAN	426-8427

Final Report

Computational and Statistical Foundations of Radar Signal Processing in the Low False Alarm Regime

AFOSR Grant No. FA9550-18-1-0190

Douglas Cochran, Arizona State University

28 July 2022

Part I
Overview

Over its four-year duration, this project has addressed two foundational problems in statistical signal processing in the context of multi-channel radar. Multistatic radar architectures are relevant in several emerging operational concepts, including purely passive localization and classification of emitters, exploitation of opportunistic sources to capitalize on energy scattered from non-emitting objects, and in networked active systems with multiple receivers and one or more transmitters.

1 Detection in Multi-channel Radar

In the initial two years, our focus was on *detection* of signals using data collected at two or more receivers. Earlier work of the PI and several other researchers, starting with [1], derived detection statistics for such scenarios that are optimal in various senses and established that these generally take the form of functions of the singular values of matrices formed from the data collected at the receivers. This project has brought computational tractability and fidelity to the evaluation of probability distributions of an important subclass of such statistics for multi-channel signal detection in the low false alarm regime. More specifically, most optimal statistics, in both Bayesian and Neyman-Pearson senses, for detection of a common signal component in $M \geq 2$ receiver channels of additive Gaussian noise are functions of the eigen-spectrum of the Gram matrix formed from segments of data collected at the sensors. When the noise is additive, complex Gaussian, and independent across the channels (i.e., under the usual H_0 hypothesis when no common signal is present), the Gram matrix is realized from a complex Wishart distribution.

Prior to the outset of this project, it had been observed that the probability distributions of desired functions of Wishart matrix eigenvalues, although known analytically, were not amenable to direct numerical evaluation – particularly for:

1. Values of the statistic corresponding to low false-alarm probabilities (e.g., $< 10^{-9}$) in signal detection scenarios,

2. Operationally significant numbers of receiver channels and signal vector dimensionality.

As discussed in more detail in Section 4 of this report, we developed a suitable transformation of the distribution of the largest eigenvalue of a central Wishart distribution (i.e., H_0 setting) that circumvents the key obstructions to direct numerical calculations when the data dimensionality is

large and the number of channels is in a realistic range. This formulation was used as a basis for calculating detection thresholds corresponding to desired false alarm rates on the order of 10^{-10} or below. In fact, the form of this algorithm enables such thresholds to be computed to essentially the limit machine of representation accuracy for problems involving several receivers and data vectors of dimension 10^5 and higher.

We subsequently extended this general approach to accurately calculate the distributions of relevant functions of the eigenvalues of non-central Wishart matrices that arise with signal present (i.e., under H_1), thus enabling full theoretical derivation of receiver operating characteristic (ROC) curves for multi-channel detectors. We were able to formulate our treatment of the H_1 problem for a rank-one signal as a rank-one perturbation of the H_0 solution and to develop suitable numerical algorithms for evaluation of correct detection probability (P_{d}) for given signal-to-noise ratio as well as false alarm probabilities (P_{f}). Some of these results are shown in Section 4 of this report.

We note that the results just mentioned, and presented more fully in [2, 3, 4], form a comprehensive solution to known numerical issues in multi-channel detection of a rank-one signal in additive Gaussian noise in the low false alarm regime. While detection of rank-one signals is the most important special case of the general problem, we also sought to address the situation where signals are of (known) rank $K = 2$ or higher (e.g., scattered signals that arise as superpositions of two or more linearly independent illuminating signals). To date, we have not been successful in addressing this case: even the simplest rank-two detection statistic (i.e., the sum of the two largest eigenvalues of a central Wishart matrix) yields marginalization integrals involving formidable combinations of special functions.

2 Signal Subspace Estimation and Tracking

The second main trust of this project also focused on statistical signal processing for multistatic radar. In this thrust, the emphasis shifted from *detecting* an unknown signal to *estimating* the subspace it defines and further to iteratively estimating that subspace as it evolves according to a dynamical model. Our approach, described in more detail in Section 5 below, yields an algorithm that mirrors the Kalman filter in this setting to the extent feasible within the mathematical constraints of the problem.

Our approach to subspace estimation and tracking is Bayesian. Consequently it relies on identifying suitable prior probability distributions for K-dimensional subspaces of an N-dimensional vector space. As discussed further in Section 5, the set of all such subspaces forms a compact manifold in which each point represents a subspace, thus the desired distributions for Bayesian estimation and tracking of subspaces will be distributions on these Grassmannian manifolds. Our results include explicit coordinatization of complex Grassmannians (the real case was done by James in 1954, but it is the complex case that is important in radar signal processing) and, more importantly for estimation, deriving explicit expressions for maximum-entropy distributions on a subclass of these manifolds corresponding to subspaces of dimension one. As demonstrated in Section 5, these distributions enable Bayesian estimation and tracking of one-dimensional subspaces of an N-dimensional vector space from data collected at multiple receivers,

In subspace estimation, as in the detection problems described above, we have not yet succeeded in extending our results from one-dimensional to higher-dimensional signals.

3 Training, Collaboration, and Dissemination

Our collaborative and dissemination activities, particularly in years two through four of the project, were negatively affected by the global Covid-19 epidemic. Nevertheless, we have continued to advance our research in multi-channel detection and have been able to maintain limited but beneficial interactions with AFRL/RY, the Australian Defence Science and Technology Group, and collaborators at Colorado State University and Dartmouth College. All papers published under support of AFOSR on multi-channel detection and subspace estimation are available from the PI's website, `http://cochran.faculty.asu.edu`.

This report concludes with a summary of our ongoing work with AFRL to place summer interns in the RY summer internship program (which continued to run remotely during the peak Covid years and is running on-site this summer) and to cultivate strong candidates among ASU's mathematics and electrical engineering students for potential careers at AFRL, other defense laboratories, and defense industry.

Part II
Technical Report

4 Multichannel Detector Numerics in the Low False Alarm Regime

This project is concerned with target detection using signal data collected at $M \geq 2$ geographically distributed receivers, as depicted in Figure 1. The measurements collected at each sensor, which in the radar setting would be antennas, are segments of narrowband waveforms. These are appropriately demodulated and sampled to form complex time series, one for each sensor. The time series are then aligned in time delay and Doppler to correspond to a putative target state (i.e., position and velocity) and a statistical test is performed to ascertain the presence of an emitter with this state whose signal impinges on the M sensors. In passive radar scenarios, the emitter will in fact be a passive scatterer of opportunistic illumination. With unknown emitter state, many feasible combinations of time delays and Doppler shifts must typically be tested, so a statistical test providing a very low false alarm rate is desirable.

If the signal source is comprised of K linearly independent emitted/scattered signals as depicted in Figure 2 (known as a "rank-K signal"), this information can be exploited in the statistical test, provided $M > K$.

In the class of multi-channel detectors that motivates this project, the M receivers are assumed to be at known positions with known velocities; the position and velocity of the putative emitter/scatterer are unknown. Also, no clutter is modeled (including direct-path signals if the illumination is opportunistic). As noted above, receivers $m = 1, \ldots, M$ each record N complex samples of received data, suitably compensated in time delay and Doppler to correspond to a target in a particular position-velocity state. These M data vectors form the rows of a data matrix X, modeled under the respective hypotheses as

$$\mathcal{H}_0 : X \sim \mathcal{CN}(0, \sigma^2 I)$$
$$\mathcal{H}_1 : X \sim \mathcal{CN}(SA, \sigma^2 I)$$

In this model, $S \in \mathbb{C}^{N \times K}$ defines the K-dimensional subspace containing the signal of interest and $A \in \mathbb{C}^{K \times M}$ contains the complex scattering coefficients.

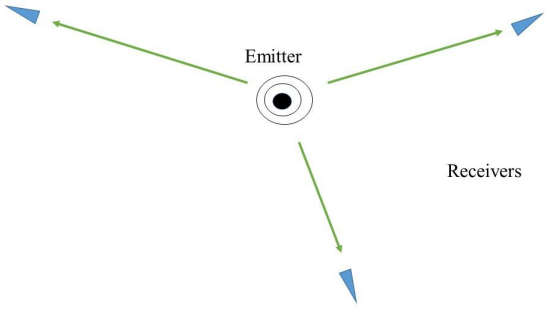

Figure 1: A notional multi-channel detection scenario.

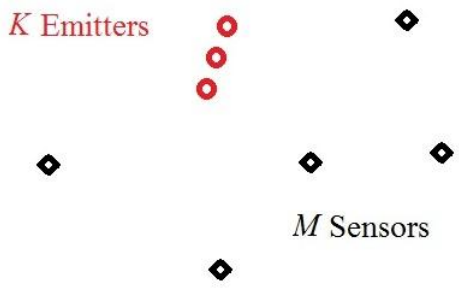

Figure 2: A notional multi-channel detection scenario with rank-K signal.

Denoting the eigenvalues of the Gram matrix $W = XX^\dagger$ as $\lambda_1 \geq \cdots \geq \lambda_M \geq 0$, detection statistics for the above problem (with known K) are [5]:

- Generalized likelihood ratio test (GLRT) statistic with known noise variance

$$\ell(X) = \sum_{i=1}^{K} \lambda_i$$

- GLRT statistic with unknown noise variance

$$\ell(X) = 1 - \frac{\sum_{i=1}^{K} \lambda_i}{\sum_{i=1}^{M} \lambda_i}$$

- Bayesian test statistics with minimally informative priors with both known and unknown noise variance are (more complicated) functions of the eigenvalues of W (see [6] and the references therein).

4.1 Analysis under H_0: Detection Thresholds

For each given test, analytically determining detection thresholds corresponding to desired false alarm probabilities requires knowledge of the conditional probability distribution of the test statistic under H_0. Under H_0, W is a complex Wishart matrix for which the joint distribution of the eigenvalues is known in closed form [7]. Unfortunately, these closed-form expressions are numerically intractable for M and N in suitable ranges for most multistatic radar applications.

The results of this project address the important special case of detecting a rank-one signal in M noisy channels, for which the GLRT statistic is the largest eigenvalue λ_1 of W. Under H_0, the distribution of λ_1 is given by Khatri [7] as

$$F_N(x) = P(\lambda_1 \leq x) = \frac{|\gamma(N+i+j, x)|_{i,j=0,\ldots,M-1}}{|\Gamma(N+i+j)|_{i,j=0,\ldots,M-1}} \tag{1}$$

with $\gamma(a, x) = \int_0^x t^{a-1} e^{-t}\, dt$. Note that the numerator and denominator in this expression are both determinants of $M \times M$ matrices whose elements are gamma functions (incomplete gamma functions in the numerator). Although the ratio of these quantities is a probability, and hence between zero and one, the factorial natures of the matrix elements quickly overwhelm floating point computation for even moderately large values of M and N. Kang and Alouini [8] give maximum values of N compatible with standard floating point computation as a function of M as tabulated in Table 1.

Sensors (M)	Maximum Samples (N)
2	98
3	71
4	57
5	47

Table 1: Largest vector length N for which expression (1) is amenable to direct computation in floating point arithmetic as a function of the number of sensors M. From [8].

As noted above, envisioned multistatic radar applications entail up to $M \approx 10$ receivers. With weak signals, time-bandwidth products $N = 10^5$ or larger are desired. Also, as previously mentioned, a large number of tests over various target states are typically required; individual tests must operate in a low false-alarm regime (e.g., $P_{\text{fa}} \approx 10^{-9}$).

Accurate Monte Carlo determination of these tail probabilities is prohibitive. In this project, the use of importance sampling and other large deviations methods for rare event simulation has been effective in extending the scope of Monte Carlo methods, but not nearly to the extent necessary to support timely calculation of accurate false-alarm probabilities in the regime of interest. Further, standard approximations (e.g., Tracy-Widom [9]) lack sufficient fidelity for this application; asymptotic techniques from random matrix theory do not apply (M and N must *both* approach infinity in a fixed ratio).

4.1.1 Transformation with Laguerre Polynomials

Based on the requirements and limitations of standard approaches just outlined, the approach pursued in this project is to transform the exact expression (1) above into a numerically useful form. Specifically, we note that the normalized (generalized) Laguerre polynomials $L_i^{(a)}(t)$ with

$a = N - 1$ are orthogonal with respect to the measure $t^{a-1}e^{-t}\,dt$; i.e.,

$$\int_0^\infty L_i^{(a)}(t)L_j^{(a)}(t)t^{a-1}e^{-t}dt = \delta_{ij}.$$

Further, they satisfy the recursion relation

$$L_{k+1}^{(a)}(t) = \frac{(2k+1+a-t)L_k^{(a)}(t) - (k+a)L_{k-1}^{(a)}(t)}{k+1}.$$

The approach proceeds to utilize orthogonality properties of $L_i^{(a)}(t)$ to eliminate gamma functions in (1) and rewrite $F_N(x)$ as inner products of polynomials.

First, we pre- and post-multiply the matrices in (1) by matrices of generalized Laguerre polynomial coefficients to obtain

$$F_N(x) = \frac{\left|\sqrt{\frac{i!j!}{(a+i)!(a+j)!}}\int_0^x L_i^{(a)}(t)L_j^{(a)}(t)t^a e^{-t}dt\right|}{\left|\sqrt{\frac{i!j!}{(a+i)!(a+j)!}}\int_0^\infty L_i^{(a)}(t)L_j^{(a)}(t)t^a e^{-t}dt\right|} = \left|\Psi^{(a)}\right|.$$

The change of variables $t \mapsto a + x\sqrt{2a}$ then yields

$$\psi_{ij}^{(a)}(a+x\sqrt{2a}) = \sqrt{2a}\sqrt{\frac{i!j!}{(a+i)!(a+j)!}} \times$$
$$\int_{-\sqrt{\frac{a}{2}}}^x L_i^{(a)}(a+t\sqrt{2a})L_j^{(a)}(a+t\sqrt{2a})(a+t\sqrt{2a})^a e^{-(a+t\sqrt{2a})}\,dt.$$

This expression for the matrix elements still has large factorials in the leading constant (recall that $a = N-1$ with N large). But this constant can be rewritten as

$$\frac{1}{\sqrt{\pi}}a^{-i/2}a^{-j/2}\frac{e^{-\epsilon(a)}}{\sqrt{\Pi_{k=1}^i(1+k/a)\Pi_{k=1}^j(1+k/a)}}$$

with

$$\begin{aligned}\epsilon(a) &= \log a! - a\log a + a - \frac{1}{2}\log 2\pi a \\ &= \frac{1}{12a} - \frac{1}{360a^3} + \frac{1}{1260a^5} - \frac{1}{1680a^7} + \dots.\end{aligned}$$

Define the functions

$$c_{ij}(a) = \frac{e^{-\epsilon(a)}}{\sqrt{\Pi_{k=1}^i(1+k/a)\Pi_{k=1}^j(1+k/a)}}$$
$$\phi_a(t) = a\log(1+t\sqrt{2/a}) - t\sqrt{2a}.$$

Using these, the integral can be rewritten as

$$\begin{aligned}&\Psi_{ij}^{(a)}(a+\sqrt{2a}x) = c_{ij}(a)\frac{\sqrt{i!j!}}{\sqrt{\pi}} \times \\ &\int_{-\sqrt{a/2}}^x a^{-i/2}L_i^{(a)}(a+t\sqrt{2a})a^{-j/2}L_j^{(a)}(a+t\sqrt{2a})e^{\phi_a(t)}dt.\end{aligned} \tag{2}$$

Note that expression (2) has no factorials of a, so the desired numerical tractability is essentially achieved at this point. It is still possible to effectively exploit the asymptotic behavior of Laguerre polynomials as follows. Define polynomials

$$D_n^{(a)}(t) = (-1)^n n!(2/a)^{n/2} L_n^{(a)}(a + t\sqrt{2a}).$$

The Laguerre and Hermite polynomials are related by [10]

$$\lim_{a\to\infty} a^{-n/2} L_n^{(a)}(a + t\sqrt{a}) = \frac{(-1)^n}{n!} 2^{-n/2} H_n(t/\sqrt{2}).$$

Therefore, by taking the limit in the number of samples N,

$$\lim_{a\to\infty} D_n^{(a)} = H_n.$$

Further, the exact distribution can be computed as the determinant of the matrix of partial inner products of the D polynomials

$$F_{a+1}(a + \sqrt{2a}x) = \left| \frac{c_{ij}(a)}{\sqrt{\pi 2^{i+j} i!j!}} \int_{-\sqrt{a/2}}^{x} D_i(t) D_j(t) e^{\phi_a(t)} dt \right|.$$

Note that the only factorials in this equation are for terms that are small in envisioned radar applications.

Finally, it is possible to capitalize upon recurrence relations for the Hermite polynomials to compute the matrix elements efficiently. This is done as follows. The above arguments show that, as $a \to \infty$, the distribution for the largest eigenvalue is

$$F_{a+1}(a + \sqrt{2a}x) = \left| \frac{1}{\sqrt{\pi 2^{i+j} i!j!}} \int_{-\infty}^{x} H_i(t) H_j(t) e^{-t^2} dt \right|.$$

Using the recurrence relations,

$$\int_{-\infty}^{x} H_i(t) H_j(t) e^{-t^2} dt = \begin{cases} \sqrt{\pi}(1 + \mathrm{erf}(x))/2 & \text{if } j = 0 \\ -H_{j-1}(x) e^{-x^2} & \text{otherwise} \end{cases}$$

so that

$$\int_{-\infty}^{x} H_i(t) H_j(t) e^{-t^2} dt = -H_i(x) H_{j-1}(x) e^{-x^2} + 2i \int_{-\infty}^{x} H_{i-1}(t) H_{j-1} e^{-t^2}\, dt.$$

4.1.2 Numerical Complementary CDF Computation

The approximate limit of problem size for computation with the 2015 method from [11] is $M = 2$ sensors and $N = 10^4$, for which the complementary cumulative distribution function is shown in Figure 3.

The approach developed enables CCDF computations for much larger problems, such as $M = 5$ and $N = 10^6$ shown in Figure 4.

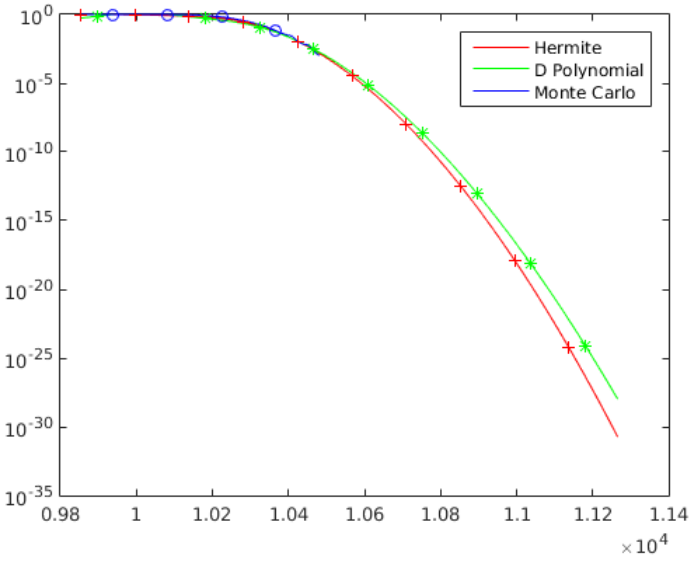

Figure 3: Complementary CDF comparison for $M = 2$ and $N = 10^4$.

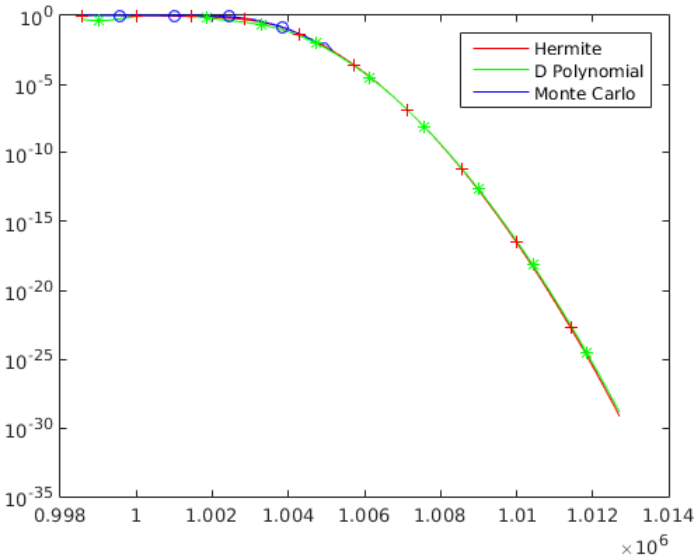

Figure 4: Complementary CDF comparison for $M = 5$ and $N = 10^6$.

4.2 H_1 Analysis: Theoretical ROC Curves

The standard of detector performance in this setting is the receiver operating characteristic (ROC) curve, which is a locus of probability of correct detection (P_d) as a function of probability of false alarm (P_f) achievable by the detector for a given signal-to-noise ratio (SNR). Note that P_f is determined from the conditional distribution of the detection statistic when no signal is present. For the GLRT detector of a rank-one signal, the detection statistic is λ_1 and thus $P_\mathrm{d} = P(\lambda_1 > T|H_0)$. Accurate computation of conditional probabilities that λ_1 exceeds a threshold value was the focus of the work described in Section 4.1. Recall that the conditional distribution of λ_1 under H_0 was obtained by Khatri and is given in Equation 1 above. The factorial natures of the matrix elements in this expression quickly overwhelm floating point computation for even moderately large values of M and N. Our results in the H_0 case, summarized above, enabled computation of the values of this distribution to arbitrary accuracy, thereby enabling detection thresholds T corresponding to desired false-alarm probabilities to be determined for ranges of M, N, and P_f encountered in real-world radar scenarios [2]. Obtaining theoretical ROC curves requires treatment of a suitable alternative hypothesis H_1 in addition to H_0.

4.2.1 Analysis Under Alternative Hypothesis

The rank-one problem is summarized as

$$\begin{aligned} H_0 &: X_m = \nu_m \\ H_1 &: X_m = \mu_m S + \nu_m \end{aligned}$$

for $m = 0, ..., M-1$, with $X_m \in \mathbb{C}^N$, the noise vectors $\nu_m \sim \mathcal{CN}(0, I)$ independent, $S \in \mathbb{C}^N$ with $||S|| = 1$ (i.e., the unit vector defining the one-dimensional signal subspace), and $\mu = (\mu_0, ..., \mu_{M-1})$. The non-zero eigenvalue of $(S\mu)(S\mu)^\dagger$ is $\mu^2 = ||\mu||^2$.

The GLRT statistic for this rank-one problem is λ_1, the largest eigenvalue of $W = XX^\dagger$. Under H_1, W is *non-central* complex Wishart distributed; the distribution of λ_1 was given in closed form by Jin *et al.* in 2008 [12] as

$$F_{\lambda_1}(x|H_1) = P(\lambda_1 \leq x|H_1) = \frac{|\Psi(x)|}{|\Psi(\infty)|}$$

where $\Psi(x)$ is the $M \times M$ matrix

$$\Psi(x) = \left[\int_0^x f_{i,j}(t)\, dt \right] \tag{3}$$

with

$$\begin{aligned} f_{i,0}(t) &= {}_0F_1(N - M + 1, \mu^2 t) t^{N-M+i} e^{-t} \qquad i = 0, ..., M-1 \\ f_{i,j}(t) &= t^{N-M+i+j} e^{-t} \qquad i = 0, ..., M-1;\ j = 1, ..., M-1. \end{aligned}$$

As in the H_0 setting, for even moderately large values of M and N, the factorial natures of the numerator and denominator quickly overwhelm floating point computation. In the following, this expression is transformed into a form suitable for tractable numerical evaluation.

4.2.2 Rank-one Perturbation on Laguerre Polynomial Transformation Method

Denote by Λ the lower triangular matrix of coefficients of the Laguerre polynomials $L_i^{(a)}(t)$ with $a = N - M$ and define ℓ to be the first column of Λ; i.e.,

$$\ell_j = \sqrt{\frac{j!}{(a-j)!}}\, L_j^{(a)}(0)$$

Conjugation of (3) by Λ yields

$$\Lambda \Psi(x) \Lambda^{\mathsf{T}} = \Phi(x) + v(x)\ell^{\mathsf{T}}$$

where

$$\Phi_{i,j}(x) = \sqrt{\frac{i!j!}{(a+i)!(a+j)!}} \int_0^x L_i^{(a)}(t) L_j^{(a)}(t) t^a e^{-t}\, dt$$

and

$$v(x)_j = \sqrt{\frac{j!}{(a-j)!}} \int_0^x [\ {}_0F_1(a+1, \mu^2 t) - 1]\ L_j^{(a)}(t) t^a e^{-t}\, dt$$

where ${}_0F_1$ denotes the confluent hypergeometric function.

Note that $\Phi(\infty) = I$, the identity matrix. Hence the matrix inversion lemma gives

$$F_{\lambda_1}(x|H_1) = \frac{|\Psi(x)|}{|\Psi(\infty)|} = \frac{(1 + \ell^{\mathsf{T}} \Phi(x)^{-1} v(x))}{(1 + \ell^{\mathsf{T}} v(\infty))}\ |\Phi(x)|$$

From prior results on the H_0 case, $F_{\lambda_1}(x|H_0) = |\Phi(x)|$; the algorithm previously derived (cf. Section 4.1.1) for computing the conditional distribution of the detection statistic λ_1 under the null hypothesis applies to this factor. So,

$$F_{\lambda_1}(x|H_1) = \frac{(1 + \ell^{\mathsf{T}} \Phi(x)^{-1} v(x))}{(1 + \ell^{\mathsf{T}} v(\infty))}\ F_{\lambda_1}(x|H_0)$$

In other words, the distribution of λ_1 under the alternative hypothesis is obtained via rank-one perturbation of the null hypothesis case. As a sanity check, note that

$$\left\{\mu^2 \to 0,\ \ {}_0F_1(a+1, 0) = 1,\ v(x) \to 0\right\} \Longrightarrow \left\{F_{\lambda_1}(x|H_1) \to F_{\lambda_1}(x|H_0)\right\}$$

4.2.3 Detection Performance Assessment

Using the above results for computing the distribution of the GLRT detection statistic λ_1 for rank-one signals under the signal-present (H_1) hypothesis together with our results for computing the distribution λ_1 under H_0), full ROC curves for this class of detector, including the low false alarm regime, can now be obtained analytically.

Figure 5 shows an example for a four-sensor scenario with a digitally modulated signal in noise at -15dB on each channel. Note that the ROC curve, which was obtained analytically, extends to $P_{\mathrm{f}} = 10^{-9}$ – well below what is tractable with Monte Carlo simulation. The data vectors are of length $N = 1000$, which is far beyond the feasible range for direct computation using (1).

Figure 6 is for a similar four-sensor scenario with vector lengths $N = 10^5$. The solid ROC curve was obtained from Monte Carlo simulation. Note that the P_{f} axis is linear in this plot; obtaining a Monte Carlo curve that extends into the low false-alarm regime as in Figure 5 is not computationally feasible. The cross symbols indicate ROC points obtained analytically. Note the close match of these points to the Monte Carlo curves across this range of P_{f} and P_{d} values.

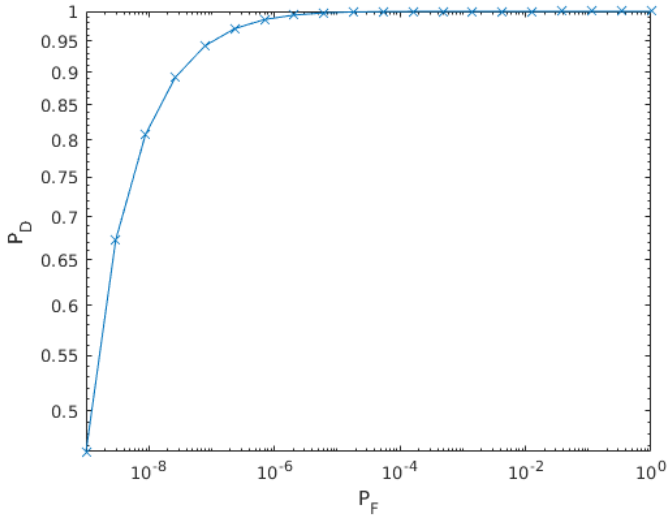

Figure 5: ROC curve $M = 4$, $N = 1000$, QPSK signal with SNR of -15dB on each channel

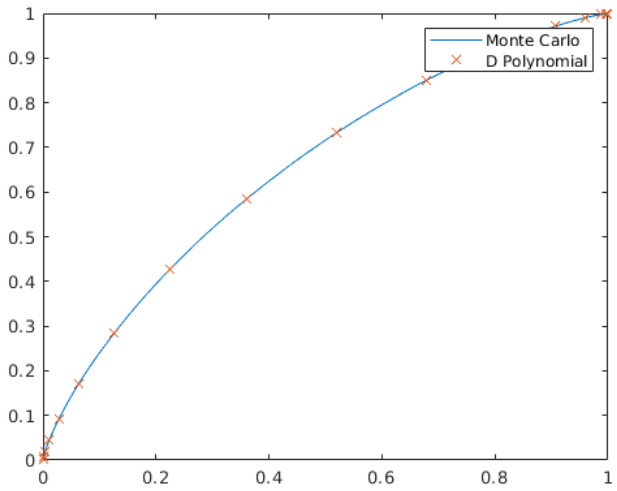

Figure 6: ROC curve $M = 4$, $N = 10^5$, QPSK signal with SNR of -15dB on each channel. Blue curve: Monte Carlo with 10^6 trials. Red crosses: Theoretical.

5 Signal Subspace Estimation

Among the most widely used signal processing algorithms in radar and telecommunications is the matched filter. In this algorithm, a known ("transmitted") signal N-vector s is to be detected in an N-vector x of received data. The datum x is, under H_0, a realization of zero-mean white Gaussian noise; under H_1, x is the sum of s with such noise. Viewed geometrically, the optimal statistic for deciding between H_0 and H_1 using data x is the magnitude of the projection of x into the one-dimensional subspace span$\{s\}$ in the ambient N-dimensional vector space. In this problem,

and many other statistical signal processing situations, it is the *subspace* defined by the signal of interest rather than the particular signal that is the essential object for such objectives as detection and estimation [14].

With a single receiver, collected data are not informative regarding the subspace defined by the transmitted/scattered signal of interest. Signal subspace models have been used effectively for many years in array signal processing where multiple sensor elements are configured in precise geometrical arrangements designed for, e.g., beamforming, direction-of-arrival estimation, or spatial filtering [15]. But deducing and exploiting signal subspace information in distributed sensor systems of the type associated with multistatic radar is a topic of recent interest and research activity [16, 17, 18, 19, 20, 21, 22, 23, 24, 5, 6].

In multi-channel detection and estimation problems, knowledge about higher-dimensional signal subspaces has been shown to enhance performance:

- If the signal subspace is is known *a priori*, it may be exploited directly [14].
- If only its dimension is known, the subspace may be estimated from collected data [16, 17, 18, 19, 20, 21].
- The dimension, if unknown, may also be estimated from data [22, 23].

5.1 Exploiting Signal Subspaces

As noted above, this project is motivated by various emerging multi-sensor remote sensing applications wherein a primary goal is to estimate a common unknown signal of known rank K measured by M spatially distributed sensors. The measurement collected at each sensor has the form of a complex N-vector, where $N \gg M$, and is regarded as an element of an ambient N-dimensional vector space V. When the number of sensors exceeds the rank of the signal ($M > K$), the rank-K signal defines a particular K-dimensional subspace, called the signal subspace, which is typically the object of interest in signal detection and estimation applications.

Consider a scenario in which a signal source emits K linearly independent waveforms in a time interval and frequency band of interest. This situation can arise from the source originating and transmitting the K signals, or it may result from a single object scattering K linearly independent waveforms impinging upon it from different directions, for example. In the absence of noise, each of the M receivers collects a different linear combination of the illuminating signals.

5.1.1 Grassmannians and Projective Space

To enable statistical analysis in this setting, particularly Bayesian analysis, it is necessary to define probability distributions on collections of subspaces of an ambient vector space V. With $K \leq N$, the collection of all K-dimensional subspaces of V forms the Grassmannian $G(K, N)$. Grassmannians are well-studied geometrical objects with several properties that are desirable from the point of view of statistical modeling. First, $G(K, N)$ is a compact Riemannian manifold of dimension $K(N-K)$. Moreover, and especially important in carrying out explicit probabilistic calculations, it is covered, except for a set of zero measure with respect to the measure described by its standard volume form, by one coordinate chart. Thus integrals over $G(K, N)$, or measurable subsets thereof, can be calculated by integrating over a single chart.

$G(K, N)$ is the natural setting for treating problems involving probabilities of collections of fixed-dimensional subspaces of a signal space V. Specifically, there is a well-defined one-to-one correspondence between K-dimensional subspaces of V and points on $G(K, N)$. Consequently,

choosing a K-dimensional subspace of V "at random" supposes a probability law on $G(K,N)$. Our work has developed such probability laws that are compatible with Bayesian estimation formulations based on Jaynes' principle of maximum entropy [25], and we have developed coordinatizations for Grassmannians that enable integrals on $G(K,N)$ that describe the probabilities of collections of K-dimensional subspaces of V to be computed by standard techniques of integration in Euclidean space.

We initially focused on the important special case of rank-one signals where V is a real vector space. In this case, $G(1,N)$ is called projective space and denoted $\mathbb{RP}^{N-1}$. In subsequent work, we developed suitable coordinatizations for complex Grassmannian manifolds; i.e., where V is a complex vector space, as is typically the case in telecommunications, radar, and electronic surveillance which involve narrowband signals on carrier waves.

5.1.2 Probabilistic Framework

In statistical signal processing, calculation of probabilities of events as well as moments and other statistical quantities generally involves calculating integrals with respect to probability measures. If x is a random variable on a space S that has probability density function f_x, then

$$P(\{x \in A\}) = \int_A f_x(y)\, dy. \tag{4}$$

When S is a (real) smooth d-manifold, this formula is valid but deceptively simple: to calculate the value of such an integral, an atlas of coordinate charts must be chosen on S and dy represents a family of differential one-forms. Ultimately, an integral is evaluated on a subset of $\mathbb{R}^d$ for each of a suitable collection of charts using ordinary calculus. These local results are appropriately stitched together to obtain the value of the global integral (4).

5.1.3 Integration on $G(K,N)$

Undertaking statistical calculations on $S = G(K,N)$ first requires identification a suitable atlas of coordinate charts. Fortunately, as discussed below, it is possible to cover the entire manifold with a single coordinate chart, except for one point which turns out to be inconsequential. We illustrate a suitable coordinatization here for the special case of real projective N-space; i.e., $\mathbb{RP}^N = G(1,N+1)$.

An element $x \in \mathbb{RP}^N$ is a one-dimensional subspace of $\mathbb{R}^{N+1}$. Let x_z be a unit vector that spans this subspace and is of the form

$$x_z = \begin{bmatrix} 1 \\ z \end{bmatrix} \left(1 + z^{\mathsf{T}} z\right)^{-1/2}, \tag{5}$$

where $z = \begin{bmatrix} z^1 & z^2 & \cdots & z^N \end{bmatrix}^{\mathsf{T}}$ is an element of $\mathbb{R}^N$. It is important to note that this representation – which we identify with the point $x \in \mathbb{RP}^N$ – is an $(N+1)$-dimensional vector whose first coordinate is positive and that these coordinates implicitly characterize a differentiable atlas comprising a single chart covering almost all of $\mathbb{RP}^n$: the set of excluded points has measure zero in the invariant measure described below. There is a corresponding orthonormal basis for the orthogonal complement of x_z in $\mathbb{R}^{N+1}$, comprising the columns of the $(N+1) \times N$ matrix which we denote $x_z^{\perp}$. In coordinates,

$$x_z^{\perp} = \begin{bmatrix} -x^{\mathsf{T}} \\ I_N \end{bmatrix} \left(I_N + z z^{\mathsf{T}}\right)^{-1/2}. \tag{6}$$

5.1.4 Differential Form for the Invariant Measure

Observing that real projective N-space is a special case of the real Grassmann manifold enables use of James' classical derivation of the invariant measure for the real Grassmann manifold [26]. Using the notation defined in the preceding section, the invariant measure on $\mathbb{RP}^N$ can be defined for points x_z in a neighborhood of a point $x \in \mathbb{RP}^N$:

$$\omega_N(x_z) \stackrel{\triangle}{=} \bigwedge_{i=1}^{N} b_i^{\mathsf{T}} \, dx_z. \tag{7}$$

In this expression, x_z is a unit vector in the subspace defined by x; dx_z is the total derivative of x_z (i.e., the differential); and $b_1, \ldots, b_N$ are the orthonormal $(N+1)$-dimensional column vectors comprising the matrix $x_z^{\perp}$. Additional calculation ultimately yields

$$\begin{aligned}
\omega_N(x) &= \alpha^{-3N/2} \det\left(\alpha\left(I_N + zz^{\mathsf{T}}\right)^{-1/2}\right) dz^1 \wedge \cdots \wedge dz^N \\
&= \alpha^{-N/2} \det\left(I_N + zz^{\mathsf{T}}\right)^{-1/2} dz^1 \wedge \cdots \wedge dz^N \\
&= \left(1 + z^{\mathsf{T}} z\right)^{-N/2} \left(1 + z^{\mathsf{T}} z\right)^{-1/2} dz^1 \wedge \cdots \wedge dz^N \\
&= \left(1 + z^{\mathsf{T}} z\right)^{-(N+1)/2} dz^1 \wedge \cdots \wedge dz^N.
\end{aligned} \tag{8}$$

The use of this form, together with transformation to hyperspherical coordinates

$$\begin{aligned}
z^k &= \tan\theta^1 \prod_{i=1}^{k} \sin\theta^i \cos\theta^{k+1} \qquad k < N \\
z^N &= \tan\theta^1 \prod_{i=1}^{n} \sin\theta^i
\end{aligned}$$

enables explicit calculation of the (known) volume of $\mathbb{RP}^N$ using elementary calculus:

$$\begin{aligned}
\mathrm{Vol}(\mathbb{RP}^N) &= \int_{\mathbb{RP}^n} \omega_N(x_z) \\
&= \int_{\mathbb{R}^n} \left(1 + z^{\mathsf{T}} z\right)^{-(N+1)/2} dz^1 dz^2 \cdots dz^N \\
&= \int_0^{\pi} \cdots \int_0^{\pi} \int_{-\pi/2}^{\pi/2} \left| \prod_{j=1}^{n} \sin^{N-j}(\theta^j) \right| d\theta^1 d\theta^2 \cdots d\theta^N \\
&= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left|\sin^{N-1}\left(\theta^1\right)\right| d\theta^1 \prod_{j=2}^{N} \int_0^{\pi} \sin^{N-j}\left(\theta^j\right) d\theta^j \\
&= \frac{\sqrt{\pi}\Gamma\left(\frac{N}{2}\right)}{\Gamma\left(\frac{N+1}{2}\right)} \prod_{j=2}^{N} \frac{\sqrt{\pi}\Gamma\left(\frac{1+N-j}{2}\right)}{\Gamma\left(\frac{2+N-j}{2}\right)} \\
&= \frac{\pi^{N/2}\Gamma\left(\frac{N}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{N+1}{2}\right)\Gamma\left(\frac{N}{2}\right)} \\
&= \frac{\pi^{\frac{N+1}{2}}}{\Gamma\left(\frac{N+1}{2}\right)}.
\end{aligned}$$

Knowledge of the differential form for the volume measure, together with transformation to hyperspherical coordinates, provides a flexible method for calculating integrals (particularly probability integrals) on projective space.

5.2 Maximum-entropy Probability Distribution

Maximum-entropy probability distributions play a crucial role in Bayesian inference using Jaynes' formulation, in which they serve as prior probability distributions that are minimally informative. Because Grassmannian manifolds are compact, they admit proper uniform densities which are just the volume form normalized by the total volume (to satisfy the requirement that a probability density must integrate to unity over the entire space). With no additional information, the uniform distribution is maximally entropic.

For the purpose of iterative estimation in the presence of a dynamical model for state propagation, collected data propagation of previous state information through the dynamical model do provide additional prior information, often in the form of moment estimates. Using a calculus of variations argument, we initially proved the following characterization for maximum-entropy probability densities on real projective space:

Theorem: For a fixed non-negative, real number κ and a fixed point $\boldsymbol{\mu}$ in $\mathbb{RP}^N$, the continuous probability density function on $\mathbb{RP}^N$ with first moment equal to $\rho\boldsymbol{\mu} \in \mathbb{R}^{N+1}$ which maximizes entropy among all such densities (on $\mathbb{RP}^N$ with first moment $\rho\mu$) is given by

$$f(\boldsymbol{x}|\boldsymbol{\mu}, \kappa) = c_N(\kappa) e^{\kappa|\boldsymbol{\mu}^{\mathsf{T}}\boldsymbol{x}|},$$

where $c_N(\kappa)$ is chosen so that $f(\boldsymbol{x}|\boldsymbol{\mu}, \kappa)$ integrates, with respect to the volume form, to unity over the whole space and κ so that $\mathsf{E}[x] = \rho(\kappa)\boldsymbol{\mu}$.

To derive the value of the normalizing constant $c_N(\kappa)$, we use the fact that

$$\int\limits_{\mathbb{RP}^N} c_N(\kappa) e^{\kappa|\boldsymbol{\mu}^\mathsf{T}\boldsymbol{x}|} d\boldsymbol{x} = 1,$$

where $d\boldsymbol{x}$ represents the differential form for the invariant measure on $\mathbb{RP}^N$. This integral may be evaluated via the transformation $\boldsymbol{y} = Qx$, where the first component of $\boldsymbol{y}$ is equal to $\boldsymbol{\mu}^\mathsf{T}\boldsymbol{x}$ and Q is orthogonal (the first row of Q will be $\boldsymbol{\mu}^\mathsf{T}$). Choosing Q orthogonal ensures $d\boldsymbol{x} = d\boldsymbol{y}$. Rewriting this as as

$$\frac{1}{c_N(\kappa)} = \int\limits_{\mathbb{RP}^N} e^{\kappa|\boldsymbol{\mu}^\mathsf{T}x|}\, d\boldsymbol{x} \int\limits_{\mathbb{R}^N} e^{\kappa\left(1+\boldsymbol{z}^\mathsf{T}\boldsymbol{z}\right)^{-\frac{1}{2}}} \frac{d\boldsymbol{z}}{\left(1+\boldsymbol{z}^\mathsf{T}\boldsymbol{z}\right)^{(N+1)/2}}.$$

Under a second transformation to hyperspherical coordinates,

$$\begin{aligned}\frac{1}{c_N(\kappa)} &= \int\limits_0^{\pi} \cdots \int\limits_0^{\pi} \int\limits_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\kappa\cos\theta^1} \left| \prod_{j=1}^{N} \sin^{N-j}(\theta^j) \right| d\theta^1 d\theta^2 \cdots d\theta^N \\ &= 2\prod_{j=2}^{N} \int\limits_0^{\pi} \sin^{N-j}\left(\theta^j\right) d\theta^j \int\limits_0^{\pm\frac{\pi}{2}} e^{\kappa\cos\theta^1} \sin^{N-1}\left(\theta^1\right) d\theta^1.\end{aligned}$$

This is now an integral in ordinary calculus, which simplifies to

$$\frac{1}{c_N(\kappa)} = \pi^{(N+1)/2} \left(\frac{2}{\kappa}\right)^{(N-1)/2} \left[I_{\frac{N-1}{2}}(\kappa) + L_{\frac{N-1}{2}}(\kappa)\right],$$

so that

$$c_N(\kappa) = \frac{\left(\frac{\kappa}{2}\right)^{\frac{N-1}{2}}}{\pi^{\frac{N+1}{2}} \left[I_{\frac{N-1}{2}}(\kappa) + L_{\frac{N-1}{2}}(\kappa)\right]},$$

where $I_\nu(\cdot)$ is the modified Bessel function of the first kind and $L_\nu(\cdot)$ is the modified Struve function.

For application in radar (and communication systems) where the signals of interest are narrow-band waveforms modulating a carrier, the natural setting is a complex vector space. In the final year of the project, we were able to extend the above result to complex projective space, $\mathbb{CP}^N$:

Theorem: For a fixed non-negative, real number κ and a fixed point $\boldsymbol{\mu}$ in $\mathbb{CP}^N$, the continuous probability density function on $\mathbb{CP}^N$ with first moment equal to $\rho\boldsymbol{\mu}$ in $\mathbb{C}^{N+1}$ which maximizes entropy among all such densities is given by

$$f_{\boldsymbol{x}}(\boldsymbol{x}|\boldsymbol{\mu},\kappa) = c_N(\kappa) e^{\kappa|\boldsymbol{\mu}^\dagger\boldsymbol{x}|},$$

where $c_N(\kappa)$ is chosen so that $f_{\boldsymbol{x}}(\boldsymbol{x}|\boldsymbol{\mu},\kappa)$ integrates, with respect to the volume form, to unity over the whole space and κ so that $\mathsf{E}[\boldsymbol{x}] = \rho(\kappa)\boldsymbol{\mu}$.

In the complex setting, $f_{\boldsymbol{x}}(\boldsymbol{x}|\boldsymbol{\mu},\kappa)$ is parameterized by its *mean direction* $\boldsymbol{\mu} \in \mathbb{CP}^N$ and its *concentration factor* $\kappa \in \mathbb{R}^+$. Further, direct calculation gives the normalization constant explicitly as

$$c_N(\kappa) = \frac{1}{\frac{\pi^N}{N!} + \left(\frac{2}{\kappa}\right)^{N-\frac{1}{2}} (\pi)^{N+\frac{1}{2}} \left[I_{N+\frac{1}{2}}(\kappa) + L_{N+\frac{1}{2}}(\kappa)\right]}$$

These maximum-entropy densities are original outcomes of this project. They are illustrated on $\mathbb{RP}^2$, in Figure 7.

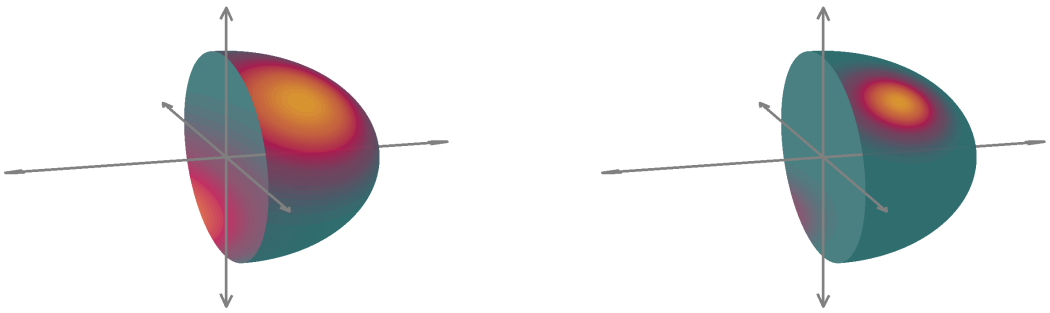

Figure 7: Two maximum-entropy probability densities on $\mathbb{RP}^2$ with the same mean direction and different concentration factors.

5.3 Iterative estimation

Recall the Kalman filter on $\mathbb{R}^d$:

1. The (post-measurement) state at time $t-1$ is a normal distribution $\mathcal{N}(M_{t-1}, \Sigma_{t-1})$.
2. This state is propagated through a linear dynamical system with additive Gaussian noise to obtain a pre-measurement state at time T; due to linearity, this state is also Gaussian.
3. A linear measurement with additive Gaussian noise is taken at time t.
4. Bayes' rule is used to produce a post-measurement (posterior) distribution from the pre-measurement state and the measurement; the linear and Gaussian assumptions imply this state is also Gaussian, $\mathcal{N}(M_t, \Sigma_t)$

From a Bayesian perspective, the Gaussian model is the maximum-entropy distribution on $\mathbb{R}^d$ with given second central moment (covariance). The linear-Gaussian dynamical and measurement models ensure everything remains Gaussian, so propagation of the state can be reduced to equations in the mean and covariance (i.e., Riccati equations).

We have reproduced key ingredients of this iterative estimation approach in the projective space setting:

- A family of maximum-entropy family of probability distributions: described in Section 5.2.
- A measurement model: starts with a standard multi-channel measurement model with Euclidean measurements in $\mathbb{C}^N$ and maps it to $\mathbb{CP}^n$ as described below.
- An invariant measure and integration method: borrowed the basics from classical work of A. T. James and recent extensions by our collaborator S. Howard; developed methods for integration.

5.4 Measurement Model

With M sensors/receivers available, M measurement vectors in $\mathbb{C}^N$ are collected in each time epoch t. At time t, the signal resides in a one-dimensional subspace of $\mathbb{C}^N$ represented by an element $\boldsymbol{s}_t$ of $\mathbb{CP}^N$. Each of the M measurements is thus a complex N-vector

$$\boldsymbol{x}_t = \boldsymbol{s}_t a_t + \boldsymbol{\nu}_t$$

where a_t defines its position in the one-dimensional subspace s_t and $\boldsymbol{\nu}_t$ is taken to be zero-mean white Gaussian noise. The update step requires that these measurements be transformed from $\mathbb{C}^N$ to elements of $\mathbb{CP}^N$, at which point estimation of the parameters of a maximum-entropy distribution on $\mathbb{CP}^N$ corresponding to the M observations is undertaken. In [13], this estimation is done using the maximum likelihood principle. The measurement density, by construction, is in the maximum-entropy class.

5.5 Iterative Algorithm

An important remark about the measurement model summarized in the preceding section is that the raw sensor measurements yield an "observed" density in the maximum entropy family. With a Bayesian approach, this is combined with the pre-measurement prior density to obtain the post-measurement density. We assume the pre-measurement density also belongs to the maximum-entropy family, as is the case with the Kalman filter. In the Kalman filter setting, the maximum-entropy densities are normal and the Bayesian combination of pre-measurement and post-measurement densities is also normal. On $\mathbb{CP}^N$, however, the Bayesian combination of two densities in the maximum-entropy family is generally not in the maximum-entropy family. Thus we implement the iterative estimator as follows:

1. At time t, pre-measurement, s_t is assumed to have a maximum-entropy distribution with parameters κ_t and $\boldsymbol{\mu}_t$.

2. Measurements are collected and a measurement distribution in the maximum-entropy family is estimated, as described above.

3. The pre-measurement and measurement distributions are combined according to Bayes' rule, yielding a distribution that is typically not in the maximum-entropy class. The post-measurement distribution is taken to be the distribution in the maximum-entropy class whose first two moments match those of the distribution produced by Bayes' rule (i.e., "moment matching").

4. A dynamical model is applied to the pre-measurement distribution for $\boldsymbol{s}_{t+1}$. If propagation by the dynamical model does not preserve the maximum-entropy form, moment matching is again used to map it into the desired family.

In our experiments, the dynamical model was a completely known unitary transformation (i.e., with no "system noise"), so the propagation in Step 4 preserves the desired form of the density and no moment matching is needed at this stage. Nevertheless, moment matching was required in Step 3. Calculation of the necessary moments was done numerically using a crude algorithm (essentially the trapezoid rule on a uniform mesh over the coordinate patch). We note than much more precise and efficient numerical integration on projective spaces and general Grassmannians is almost certainly possible.

5.6 Numerical Example

A numerical example illustrates the theory developed for the tracker on $\mathbb{RP}^4$, when the measurements collected at each epoch are noisy samples of a signal whose subspace is defined by the true state at time t. The parameters σ^2 and a_t are adjusted to control the signal-to-noise ratio (SNR) for each measurement.

The true, veiled subspace is rotated by a deterministic and known element of $\mathbf{SO}(\mathbf{5})$ and the recursive estimation algorithm described in the preceding Section is performed. At each time

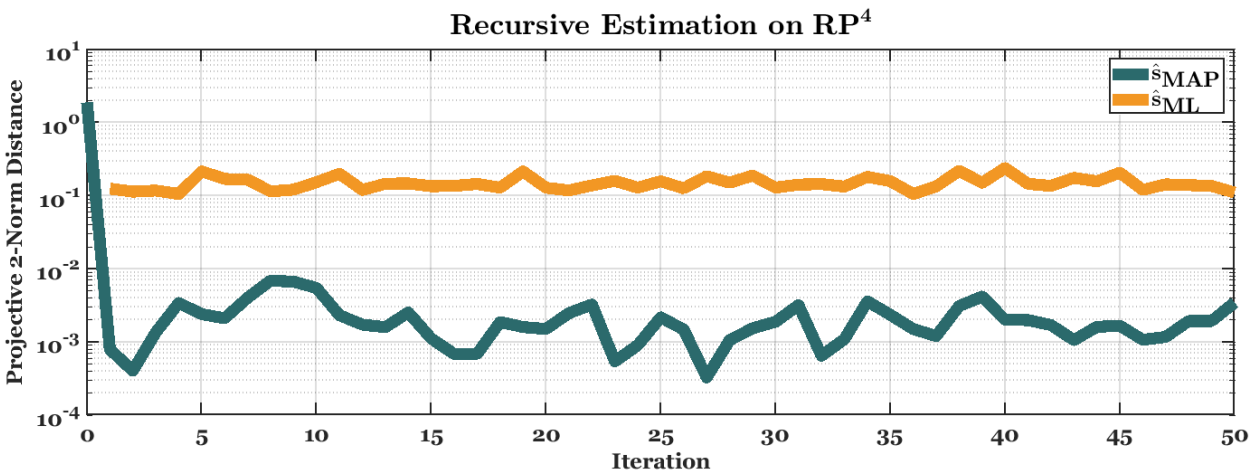

Figure 8: Tracking on $\mathbb{RP}^4$ with measurements in $\mathbb{R}^5$ that are samples of a noisy signal whose subspace is defined by an element of $\mathbb{RP}^4$.

epoch, the maximum *a posteriori* point estimate $\hat{\boldsymbol{s}}_{t,\mathrm{MAP}}$ of the unknown subspace is computed from its posterior distribution, as well as the maximum-likelihood point estimate $\hat{\boldsymbol{s}}_{t,\mathrm{ML}}$ from the sensor measurements in $\mathbb{R}^5$. Both point estimates are compared to the true subspace at time t. Plotted on a log scale in Figure 8 are the distances between the two point estimates and the truth, defined by the projection 2-norm, i.e., by the sine of the principal angle between the two subspaces. The distance between the recursive maximum *a posteriori* point estimate $\hat{\boldsymbol{s}}_{t,\mathrm{MAP}}$ and the truth decreases with each iteration of Bayesian estimation algorithm, while the distance between the maximum-likelihood point estimate $\hat{\boldsymbol{s}}_{t,\mathrm{ML}}$ remains effectively static and significantly inferior to the recursive maximum *a posteriori* point estimate. There are $m = 10$ measurements observed at each time epoch, and the SNR on each sensor is set to approximately 0 dB. We note that, while encouraging, these results are for a dynamical system with no process noise.

Part III
Transition and Outreach

Historically, our AFOSR-sponsored research activities have entailed regular interaction between the PI's research group at ASU and elements of the Sensors Directorate at AFRL, and we collaborate closely with researchers in the Cyber and Electronic Warfare Division and the National Security and ISR Division of the Australian Defence Science and Technology Group. Many of these research interactions have been three-way (i.e., AFRL-DSTG-ASU). As noted in the introduction of this report, these interactions have been seriously attenuated over the last two and a half years of the project. Nevertheless, we have continued to participate in individual technical discussions with AFRL personnel and in frequent tele-meetings organized by AFRL.

This project has been instrumental in enabling ASU's ongoing participation in the summer internship program organized by the ATR Center in AFRL/RY. We have had twenty-six graduate and undergraduate students (approximately 50% women) complete summer internships at WPAFB under this program since 2012, including several who are now in professional R&D positions in DoD laboratories or in U.S. defense industry. These experiences and career trajectories have created positive visibility of Defense and National Security career possibilities among ASU students. The ATRC summer program was run remotely in 2020 and 2021, and it has returned to an on-line format this year. We have maintained a good level of participation, both through involvement in student project mentorship and by arranging ASU students' enrollment in the program. In summer, 2022, ASU students Victoria Uribe and Nicholas Chmielewski (doctoral students in Applied Mathematics) are at WPAFB for the program.

Scott Jones was a key contributor to this project until he completed his Ph.D. studies in August 2019 [4]. He is now a research scientist at Slingshot Aerospace. Lauren Crider was the primary student contributor to this effort from 2019 until she completed her Ph.D. in Fall 2021 [13]. She has taken a career position in industry.

References

[1] D. Cochran, H. Gish, and D. Sinno, "A geometric approach to multiple-channel signal detection," *IEEE Transactions on Signal Processing*, vol. 42, no. 9, pp. 2049–2057, 1995.

[2] S. Jones, S. D. Howard, I. V. L. Clarkson, K. Bialkowski, and D. Cochran, "Computing the largest eigenvalue distribution for complex Wishart matrices," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 3439–3443, April 2017.

[3] S. Jones, S. D. Howard, I. V. L. Clarkson, K. Bialkowski, and D. Cochran, "Computing the largest eigenvalue distribution for non-central Wishart matrices," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 4290–4294, May 2019.

[4] S. Jones, *Numerical Computation of Wishart Eigenvalue Distributions for Multistatic Radar Detection*, Ph.D. dissertation, Arizona State University, July 2019.

[5] S. Sirianunpiboon, S. D. Howard, and D. Cochran, "Multiple-channel detection of signals having known rank," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 6536–6540, May 2013.

[6] S. D. Howard, S. Sirianunpiboon, and D. Cochran, "An exact Bayesian detector for multistatic passive radar," *Proceedings of the Asilomar Conference on Signals, Systems, and Computers*, pp. 1077–1080, November 2016.

[7] C. G. Khatri, "Distribution of the largest or the smallest characteristic root under null hypothesis concerning complex multivariate normal populations," *Annals of Mathematical Statistics*, vol. 35, no. 4, pp. 1807–1810, 1964.

[8] M. Kang and M. S. Alouini, "Largest eigenvalue of complex Wishart matrices and performance analysis of MIMO MRC systems," *IEEE Journal on Selected Areas in Communications*, vol. 21, No. 3, pp. 418–426, 2003.

[9] C. A. Tracy and H. Widom, "Distribution functions for largest eigenvalues and their applications," *Proceedings of the International Congress of Mathematicians*, vol. I, pp. 587–596, 2012.

[10] N. M. Temme, *Special functions: An introduction to the classical functions of mathematical physics*, Wiley, 1996.

[11] K. S. Bialkowski and I. V. L. Clarkson, "Computing multistatic passive radar CFAR thresholds from surveillance-only data," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 5555–5559, 2015.

[12] S. Jin, M. R. McKay, X. Gao, and I. B. Collings, "MIMO multichannel beamforming: SER and outage using new eigenvalue distributions of complex noncentral Wishart matrices," *IEEE Transactions on Communications*, vol. 56, no. 3, pp. 424–434, March 2008.

[13] L. N. Crider, *Recursive Bayesian estimation on projective spaces*, Ph.D. thesis, Arizona State University, 2001.

[14] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Transactions on Signal Processing*, vol. 8, pp. 2146–2157, 1994.

[15] H. Krim and M. Viberg, "Two decades of array signal processing research," *IEEE Signal Processing Letters*, vol. 13, no. 4, pp. 67–94, 1996.

[16] S. Sirianunpiboon, S. D. Howard, and D. Cochran, "Exact Bayesian test for a common rank-one component in white noise," *Proceedings of the Asilomar Conference on Signals, Systems, and Computers*, pp. 1496–1499, November 2015.

[17] D. E. Hack, L. K. Patton, and B. Himed, "Multichannel detection of an unknown rank-one signal with uncalibrated receivers," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 2987–2991, May 2014.

[18] S. Sirianunpiboon, S. D. Howard, and D. Cochran, "Multiple-channel detection of signals having known rank," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 6536–6540, May 2013.

[19] D. E. Hack, C. W. Rossler, and L. K. Patton, "Multichannel detection of an unknown rank-N signal using uncalibrated receivers," *IEEE Signal Processing Letters*, vol. 21, no. 8, pp. 998–1002, 2014.

[20] D. Ramírez, G. Vazquez-Vilar, R. López-Valcarce, J. Vía, and I. Santamaría, "Detection of rank-P signals in cognitive radio networks with uncalibrated multiple antennas," *IEEE Transactions on Signal Processing*, vol. 59. no. 8, pp. 3764–3775, 2011.

[21] D. Ramírez, J. Iscar, J. Vía, I. Santamaría, and L. L. Scharf, "The locally most powerful invariant test for detecting a rank-P Gaussian signal in white noise," *Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop*, pp. 493–496, June 2012.

[22] S. D. Howard, S. Sirianunpiboon, and D. Cochran, "Detection and characterization of MIMO radar signals," *Proceedings of the IEEE International Radar Conference*, pp. 330–334, September 2013.

[23] S. Sirianunpiboon, S. D. Howard, and D. Cochran, "Maximum a posteriori estimation of signal rank," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 5676–5680, May 2014.

[24] S. D. Howard and S. Sirianunpiboon, "Passive radar detection using multiple transmitters," *Proceedings of the Asilomar Conference on Signals, Systems, and Computers*, pp. 945–948, November 2013.

[25] E. T. Jaynes, "Information theory and statistical mechanics," *Physical Review*, vol. 106, no. 4, pp. 620–630, 1957.

[26] A. T. James, "Normal multivariate analysis and the orthogonal group," *The Annals of Mathematical Statistics*, vol. 25, no. 1, pp. 40–75, 1954.