Fast Estimation of False Alarm Probabilities
of STAP Detectors - The AMF

Rajan Srinivasan, University of Twente, Netherlands
Muralidhar Rangaswamy, Air Force Research Laboratory, USA

ABSTRACT

This paper describes an attempt to harness the power of adaptive importance sampling techniques for estimating false alarm probabilities of detectors that use space-time adaptive processing. Fast simulation using these techniques have been notably successful in the study of conventional constant false alarm rate radar detectors, and in several other applications. The principal task here is to examine the viability of using importance sampling methods for STAP detection. Though a modest beginning, the adaptive matched filter detection algorithm is analysed successfully using fast simulation. Of two biasing methods considered, one is implemented and shown to yield excellent results. The important problem of detector threshold determination is also addressed, with matching outcome. The work reported here serves to pave the way to development of more advanced estimation techniques that can facilitate design of powerful and robust detection algorithms designed to counter hostile and heterogenous clutter environments.

1. *INTRODUCTION*

Estimation of false alarm probabilities of detection algorithms that employ space-time processing is examined here using forced Monte Carlo (MC) or importance sampling (IS) simulation. Space-time adaptive processing (STAP) algorithms are of much importance for radar detection. They are notoriously intensive from a computational point of view, with the more advanced (and robust) ones being also analytically difficult to quantize, [1], [2]. Therefore it is appropriate to attempt to develop fast simulation methods that can be used in their analysis and design.

In this paper we use lessons learnt from developing adaptive IS techniques for characterizing conventional constant false alarm rate (CFAR) detectors, [3], [4], and describe an experiment in applying them to STAP detection. The starting point of this effort is the celebrated adaptive matched filter (AMF) derived in [1], and which represents the arrayed version of the workhorse cell averaging CFAR detector for conventional radar signal processing. The false alarm probability (FAP) performance of the AMF detector is known in integral form and can be numerically computed to any desired accuracy. Thus it forms a suitable basis for validating our simulation experiments. Two specific IS methods (described in the sequel) are presented and the better (and also computationally easier) one is implemented. On a general note, IS is the chief simulation methodology for rare event estimation. It is an enduring method that has distinguished itself in several areas of science and engineering, [5], [6]. Briefly, IS works by biasing original probability distributions in ways that accelerate the occurrences of rare events, conducting simulations with these new distributions, and then compensating the obtained results for the changes made. The principal consequence of this procedure is that unbiased probability estimates with low variances are obtained quickly. The main task in IS therefore is determination of good simulation distributions for an application, either as a one-shot feat or adaptively. Simulations performed using such distributions can provide enormous speed-ups if they are chosen with due care and mathematical precision. Indeed, if applied successfully, simulation lengths needed to estimate very low probabilities become (only) weakly dependent on the actual probabilities. It is thus possible to evaluate any probability in reasonable amounts of simulation time. There have been more recent attempts in the literature, for example [7], [8], to apply IS for FAP estimation of CFAR detectors with varying degrees of success. The work reported here however, relies entirely on the methods developed in [3] - [5].

During the conduct of simulations reported herein, some issues arose concerning the adaptive IS algorithms used, and these are discussed briefly. More investigation is required into them. However, the positive outcome is that excellent match with numerical results is obtained. The succeeding sections provide a short statement of the AMF algorithm, how IS biasing can be performed to hasten false alarm events, description of the so called g-method which is a conditional IS technique developed originally for studying sums of random variables ([3]), the fast algorithms used, how *inverse IS* can be used to estimate (and therefore design) detector thresholds, simulation results, and a concluding discussion.

2. *THE AMF DETECTOR*

In a radar system consisting of a linear array of N_s antenna elements, a burst of N_t pulses is transmitted and each element receives as many return samples in any one range gate. The $N_s N_t = N$ samples are complex (because of I and Q channel

processing) and are referred to as the primary data. They may contain a target and represent the range gate to be tested. The samples are arranged in an $N \times 1$ column vector and denoted as $\mathbf{x}$. The target return is modelled as consisting of a known direction vector $\mathbf{s}$ with an unknown complex amplitude in addition to clutter, interference, and noise. There are L other N-length complex vectors, called the secondary data, obtained from as many nearby range gates and assumed to be free of target signal. These are denoted as $\mathbf{x}(l)$, $l = 1, \ldots, L$. The primary and secondary data vectors are assumed to be jointly independent and complex Gaussian, sharing the $N \times N$ covariance matrix $\mathbf{R} = E\{\mathbf{X}\mathbf{X}^\dagger\}$, where the superscript $\dagger$ denotes complex conjugate transpose. Under these assumptions the AMF detection test, as obtained in [1], is given by

$$\frac{|\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1} \mathbf{x}|^2}{\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1} \mathbf{s}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \qquad (1)$$

where

$$\widehat{\mathbf{R}} \equiv \frac{1}{L} \sum_{l=1}^{L} \mathbf{x}(l)\mathbf{x}(l)^\dagger$$

is the estimated covariance matrix of $\mathbf{x}$ based on the secondary data (also referred to as sample matrix), and η is a threshold used to set the FAP at some desired level. This test has the CFAR property. The FAP α of the test is known to be given by

$$\alpha = \frac{L!}{(L-N+1)!(N-2)!} \int_0^1 \frac{x^{L-N+1}(1-x)^{N-2}}{(1+\eta x/L)^{L-N+1}} \, dx \qquad (2)$$

which can be used to numerically determine the threshold setting for a desired FAP. As shown in [1], the test in (1) can be rewritten as

$$\begin{aligned}
|\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}|^2 \quad &\underset{H_0}{\overset{H_1}{\gtrless}} \quad \eta\,\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s} \\
&= \quad \eta\,\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\widehat{\mathbf{R}}\widehat{\mathbf{R}}^{-1}\mathbf{s} \\
&= \quad \eta\,\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\frac{1}{L}\sum_{l=1}^{L}\mathbf{x}(l)\mathbf{x}(l)^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s} \\
&= \quad \frac{\eta}{L}\sum_{l=1}^{L}\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{x}(l)\mathbf{x}(l)^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s} \\
&= \quad \frac{\eta}{L}\sum_{l=1}^{L}|\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{x}(l)|^2 \qquad (3)
\end{aligned}$$

This is in the form of a vector (or, array) version of the usual CA-CFAR test. The LHS is a square law detector, being the output of a matched filter (matched to the direction $\mathbf{s}$ in which the array is steered) for incoherent detection using the so-called sample matrix inversion beamformer weights $\widehat{\mathbf{R}}^{-1}\mathbf{s}$. The RHS represents a cell averaging term. Further details on these issues can be found in the references mentioned above.

3. FALSE ALARM PROBABILITY ESTIMATION USING IS

Two methods to quickly estimate FAPs are two-dimensional (2-d) biasing and the conditional g-method

procedure, described in this section.

3.1. 2-d Biasing

We make the following observations. Suppose each complex sample of a secondary vector is scaled by a real number $\theta^{1/2}$. This has the effect of scaling the covariance matrix estimate $\widehat{\mathbf{R}}$ by θ. Therefore, as far as the covariance estimate is concerned, both sides of the test in (3) remain unaffected by the scaling. However, each secondary vector being scaled by $\theta^{1/2}$ results in a scaling of the RHS by θ. Hence choosing θ less than unity will have the effect of compressing the density function of the random threshold of the test. Further, a scaling of each complex component of the primary vector by a real $a^{1/2}$ will achieve a scaling of the LHS of the test by a. Thus, choosing a larger and θ smaller than unity will achieve an increase in the frequency of occurrence of false alarm events during simulation. The IS optimization problem will be a two-parameter one.

The (unbiased) IS estimator, using (1), can be expressed as

$$\widehat{\alpha} = \frac{1}{K}\sum_1^K 1(|\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{X}|^2 > \eta\,\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s})\,W(\mathbf{X},\mathbf{X}_L; a, \theta);$$
$$\sim f_\star \qquad (4)$$

where $1()$ is an indicator, the notation $\sim f_\star$ means that all random variables are drawn from biased distributions, and $\mathbf{X}_L \equiv (\mathbf{X}(1), \ldots, \mathbf{X}(L))^T$, K being length of the IS simulation. In setting up their joint densities, we use the fact that the FAP of the AMF has the CFAR property and is independent of the true covariance matrix $\mathbf{R}$. This is true under the assumption of Gaussian distributions for the data. In such a case, simulation of the AMF test can be carried out for data possessing an $N \times N$ identity matrix $\mathbf{I}$ as covariance. Therefore, primary and secondary data can be generated as complex vectors with independent components. The unbiased joint densities are

$$f(\mathbf{x}) = \frac{e^{-\mathbf{x}^\dagger\mathbf{x}}}{\pi^N} \text{ and } f(\mathbf{x}_L) = \frac{e^{-\sum_1^L \mathbf{x}(l)^\dagger\mathbf{x}(l)}}{\pi^{LN}}$$

so that

$$f(\mathbf{x}, \mathbf{x}_L) = \frac{e^{-\mathbf{x}^\dagger\mathbf{x}-\sum_1^L \mathbf{x}(l)^\dagger\mathbf{x}(l)}}{\pi^{(L+1)N}}$$

With scaling performed as described above, the biased joint density takes the form

$$f_\star(\mathbf{x}, \mathbf{x}_L) = \frac{e^{-\frac{1}{a}\mathbf{x}^\dagger\mathbf{x}-\frac{1}{\theta}\sum_1^L \mathbf{x}(l)^\dagger\mathbf{x}(l)}}{\pi^{(L+1)N}a^N\theta^{LN}}$$

resulting in the weighting function

$$\begin{aligned}
W(\mathbf{X}, \mathbf{X}_L; a, \theta) &\triangleq \frac{f(\mathbf{x}, \mathbf{x}_L)}{f_\star(\mathbf{x}, \mathbf{x}_L)} \\
&= Ca^N\theta^{LN}e^{A/a}e^{B/\theta} \qquad (5)
\end{aligned}$$

where

$$A \equiv \mathbf{x}^\dagger\mathbf{x}, \quad B \equiv \sum_1^L \mathbf{x}(l)^\dagger\mathbf{x}(l), \text{ and } C \equiv e^{-(A+B)}$$

The variance of the IS estimator $\widehat{\alpha}$ can be expressed as

$$\operatorname{var}\widehat{\alpha} = \frac{1}{K}[I(\boldsymbol{\nu}) - \alpha^2] \qquad (6)$$

where $\boldsymbol{\nu}$ is the vector biasing parameter $(a, \theta)^T \in [1, \infty) \times (0, 1]$. Denoting by $\mathcal{A}$ the false alarm event in (4), the quantity I is given by

$$
\begin{aligned}
I(\boldsymbol{\nu}) &= E_\star\{1(\mathcal{A})W^2(\mathbf{X}, \mathbf{X}_L; \boldsymbol{\nu})\} \\
&= E\{1(\mathcal{A})W(\mathbf{X}, \mathbf{X}_L; \boldsymbol{\nu})\} \qquad (7)
\end{aligned}
$$

where the expectation $E_\star$ proceeds over biased distributions. Minimization of $\operatorname{var}\widehat{\alpha}$ with respect to the biasing parameters is equivalent to minimization of I and is described in the Appendix. Although not implemented here, this description has been included since it is foreseen that such a method could be useful in situations wherein the g-method might be difficult to apply.

3.2. *The g-method Estimator*

This method exploits knowledge of underlying distributions more effectively, yielding a more powerful estimator. Additional advantages are that only a scalar parameter optimization problem needs to be tackled and the *inverse IS* problem (for threshold optimization) can be easily solved. The FAP can be written as

$$
\begin{aligned}
\alpha &= P(|\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{X}|^2 > \eta\,\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s}|H_0) \\
&= E\{P(|\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{X}|^2 > \eta\,\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s}|\mathbf{X}_L, H_0)\} \\
&\triangleq E\{g(\mathbf{X}_L)\} \qquad (8)
\end{aligned}
$$

Note that the conditioning in the second line above implies the condition that a covariance matrix estimate is given. We proceed to estimate α using the form in the third line above.

With the condition in mind it is easy to show, assuming that $\mathbf{X}$ is rotationally invariant and Gaussian, that the random variable $\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{X} \triangleq \mathbf{w}^\dagger\mathbf{X}$ is distributed as $\mathcal{CN}_1(0, \mathbf{w}^\dagger\mathbf{R}\mathbf{w})$ with independent real and imaginary components, and the weight vector $\mathbf{w} = \widehat{\mathbf{R}}^{-1}\mathbf{s}$. The random variable $Y \triangleq |\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{X}|^2$ therefore is exponential and has density function

$$f(y|\mathbf{X}_L, H_0) = \frac{e^{-y/\mathbf{w}^\dagger\mathbf{R}\mathbf{w}}}{\mathbf{w}^\dagger\mathbf{R}\mathbf{w}}, \quad y \geq 0$$

Therefore

$$
\begin{aligned}
g(\mathbf{X}_L) &= P(Y \geq \eta\,\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s}|\mathbf{X}_L, H_0) \\
&= e^{-\eta\,\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s}/\mathbf{w}^\dagger\mathbf{R}\mathbf{w}}
\end{aligned}
$$

Note that if $\widehat{\mathbf{R}} = \mathbf{R}$, then $g(\mathbf{X}_L) = e^{-\eta}$ and this is the FAP of the AMF when the covariance matrix is known. As discussed before, we are simulating with homogeneous data possessing an identity covariance matrix, that is, with $\mathbf{R} = \mathbf{I}$. The g-method IS estimator then takes the form

$$
\begin{aligned}
\widehat{\alpha}_g &= \frac{1}{K}\sum_1^K g(\mathbf{X}_L)W(\mathbf{X}_L; \theta) \\
&= \frac{1}{K}\sum_1^K e^{-\eta D}W(\mathbf{X}_L; \theta); \quad \sim f_\star \qquad (9)
\end{aligned}
$$

where

$$
\begin{aligned}
D &\equiv \frac{\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s}}{|\mathbf{w}|^2} \\
&= \frac{\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s}}{\mathbf{s}^\dagger(\widehat{\mathbf{R}}^{-1})^2\mathbf{s}}
\end{aligned}
$$

Choosing the (single) biasing parameter $\theta < 1$ thus produces a decrease in D, thereby causing a higher frequency of occurrence of the false alarm event or, more appropriately in this case, a larger value of the g-function. Note that use of the g-method obviates the need to bias primary data vectors. Determination of a good value of θ proceeds as before. The weighting function is simply

$$W(\mathbf{x}_L; \theta) = \theta^{LN}e^{-(1-1/\theta)B} \qquad (10)$$

which can be deduced from (5) by setting $a = 1$. The scaling θ is optimized by

$$\theta_{m+1} = \theta_m - \delta_\theta\frac{\widehat{I}_g'(\theta_m)}{\widehat{I}_g''(\theta_m)} \qquad (11)$$

which is just a one-dimensional version of (15). Estimates of the I-function and its derivatives are given by

$$
\begin{aligned}
\widehat{I}_g(\theta) &= \frac{1}{K}\sum_1^K g^2(\mathbf{X}_L)W^2(\mathbf{X}_L; \theta); \quad \sim f_\star \\
\widehat{I}_g'(\theta) &= \frac{1}{K}\sum_1^K g^2(\mathbf{X}_L)W(\mathbf{X}_L; \theta)W_\theta(\mathbf{X}_L; \theta); \quad \sim f_\star \\
\widehat{I}_g''(\theta) &= \frac{1}{K}\sum_1^K g^2(\mathbf{X}_L)W(\mathbf{X}_L; \theta)W_{\theta\theta}(\mathbf{X}_L; \theta); \quad \sim f_\star
\end{aligned}
$$

See the Appendix for definition of the above quantities.

3.2.1 *Simulation Gain*

A useful measure of the effectiveness of any IS procedure is the simulation gain Γ. It is the ratio of simulation lengths required by conventional MC and IS estimators to achieve the same error variance. Setting the variance in (6) equal to $(\alpha - \alpha^2)/k$ (being the MC variance) where k denotes the length required by the MC estimator, yields the gain

$$\Gamma = \frac{\alpha - \alpha^2}{I(\boldsymbol{\nu}) - \alpha^2}$$

While the simulation gain is useful in learning how much faster than MC an IS technique is in terms of simulation length, it also serves the purpose of comparing different IS estimators. In actual simulations, an estimate of Γ is made by using the estimates for α and I. The g-method estimator has simulation gain given by

$$\Gamma_g = \frac{\alpha - \alpha^2}{I_g(\boldsymbol{\theta}) - \alpha^2}$$

where $I_g = E_\star\{g^2(\mathbf{X}_L)W^2(\mathbf{X}_L; \theta)\}$, and it can be estimated during simulation. It always has a smaller variance and consequently larger gain than the IS estimator discussed in the previous section. Indeed, without IS ($W = 1$),
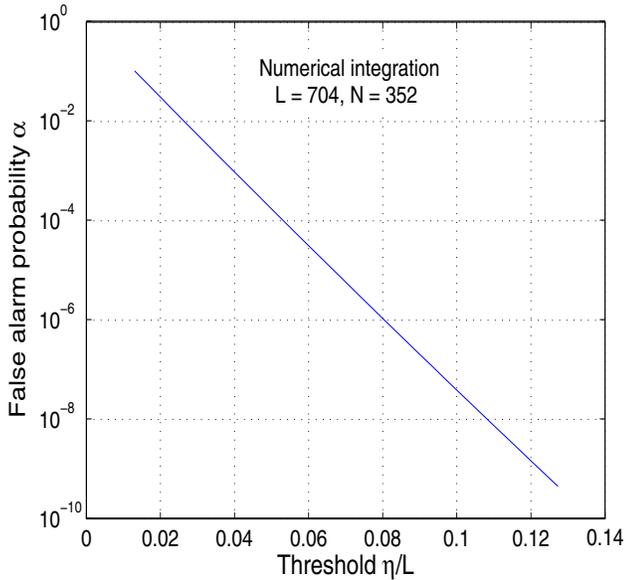
Fig. 1.   Numerically computed FAP of the AMF detector.

Fig. 2.   Threshold optimization for AMF detector using inverse IS algorithms.

$I_g = E\{g^2(\mathbf{X}_L)\} < E\{g(\mathbf{X}_L)\} = \alpha$. That $I_g < I$ with IS was proved in [3] for conventional CFAR detectors. The proof in the case of the detectors considered here is similar, and will be omitted.

3.3. *Inverse IS and Threshold Determination*

The inverse problem [3], namely, that of finding by fast simulation the value of detector threshold η satisfying a prescribed FAP, is readily solved using the g-method estimator. This is done by minimizing the stochastic objective function

$$J(\eta) = [\widehat{\alpha}_g(\eta) - \alpha_o]^2$$

where α_o is the desired FAP. Minimization of J with respect to η is carried out by the algorithm

$$\eta_{m+1} = \eta_m + \delta_\eta \frac{\alpha_o - \widehat{\alpha}_g(\eta_m)}{\widehat{\alpha'_g}(\eta_m)}, \quad m = 1, 2, \dots \quad (12)$$

where δ_η is a step-size parameter and the derivative estimator in the denominator is given by

$$\widehat{\alpha'_g}(\eta_m) = -\frac{1}{K} \sum_1^K D\, e^{-\eta D} W(\mathbf{X}_L; \theta); \quad \sim f_\star \quad (13)$$

with prime indicating derivative with respect to η. Note in passing that this derivative estimator actually estimates (negative of) the probability density function of the AMF statistic on the left hand side of (1) under H_0. The algorithms (4), (11), and (12) are implemented simultaneously.

4. *NUMERICAL RESULTS*

The FAP α obtained by direct numerical integration of (2) is shown in Fig. 1 and used for comparing IS results displayed in the remaining figures. Detector parameters are
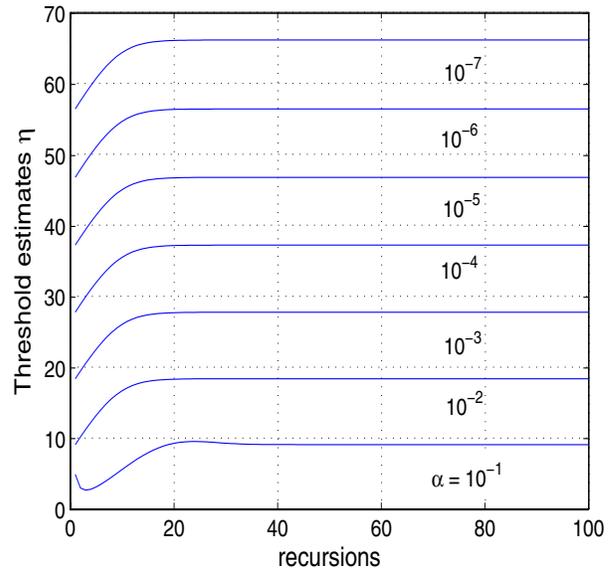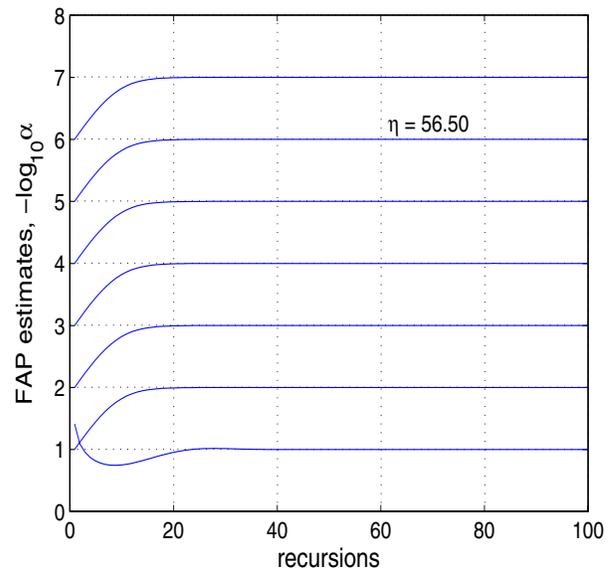

Fig. 3.   FAP estimates resulting from threshold optimization algorithms.

$L = 704$ and $N = 352$. Fig. 2 and Fig. 3 show results of implementing the inverse IS algorithms. These are estimated threshold settings for pre-specified FAPs. Optimum biasing parameters are shown in Fig. 4 and resulting simulation gains in Fig. 5. It is evident that match with Fig. 1 is excellent and this has been numerically confirmed.

4.1. *Discussion*

These IS simulation results appear smooth and certainly beg an obvious question. Indeed, an artifice has been employed here to generate them. It was used by the first author in
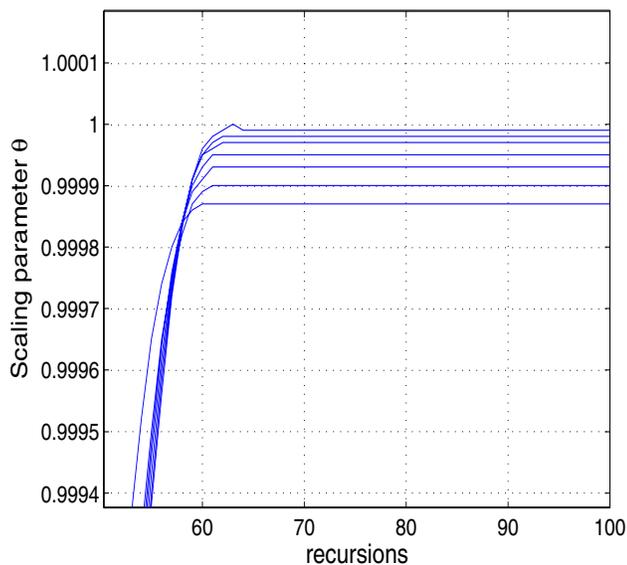
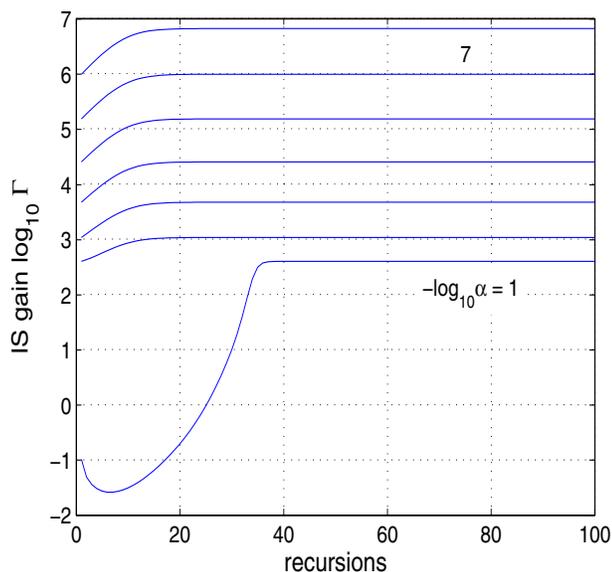Fig. 4.   Optimum biasing (scaling) parameters for IS.

Fig. 5.   Estimated simulation gains.

previous work on capacity estimation of MIMO channels and elsewhere, and found to be extremely useful. In conducting rare-event simulations of systems that involve mathematically complex signal processing operations, there are two principal issues that contribute to simulation time. These have to be dealt with effectively. The first concerns the rare event itself whose probability is being sought, and this can of course be handled by suitable IS techniques. The second (but related) issue is the computational intensity that accompanies the signal processing. In the case of STAP detectors, the chief processing burden is from inversion of large sample matrices. It is a daunting task to conduct conventional MC simulations involving several millions of trials to estimate low FAPs, with as many inversions. Assuming that a good IS scheme can

reduce the number of trials to, say, only a few thousands would still be computationally demanding (a case in point being the three thousand 352×352 matrices that were inverted here). This is where IS departs from conventional MC in a subtle yet important manner. It is almost totally useless to run the same random variables through a system in a straight MC simulation. With IS however, much can be learnt by repeatedly using the (same) random variables. In fact, this is one of the powerful features that adaptive IS (and inverse IS) can embed into complex system simulation. But how does an IS scheme become effective in the first place? Assume that we have a biasing scheme that promises to be effective once the parameters of the biasing distributions have been optimized. For large systems (in the sense of number of random inputs involved), running truly randomized IS algorithms adaptively could become demanding as pointed out above. If system performance can be characterized in terms of certain random 'metrics' (we use the word with a slight abuse of terminology), then these metrics can be pre-computed for a given set of input variables, and used repeatedly (which, in complex systems such as STAP detectors, eases the computational burden) in adaptive biasing optimization algorithms. These latter algorithms themselves often require no more than 100 iterations and can be extremely fast. Resulting IS simulation gains can be simultaneously estimated and these tell us whether we need more or less pre-generated variables to achieve certain accuracies. Adjusting this latter number, biasing and system parameter optimization (inverse IS) algorithms can be run, once. Thus there is an initial step during which gains are estimated based on pre-computed metrics and then the number of these metrics is adjusted. All this is not as complicated as it appears. Turning attention to (9), (10), and (13), the only two random quantities (or metrics) needed to estimate FAP and associated detection threshold are B and D. This is for the g-method. For 2-d biasing the only additional quantity required is the norm A of the primary vector, defined just after (5) and this adds almost nothing to the computation. It turned out that generating $K = 3000$ random instances of B and D was certainly an overkill. If one looks at Fig. 5, the gain provided by IS for estimating $\alpha = 10^{-6}$ is about 10^6. From usual asymptotic normality arguments, [3], it follows that about 100 optimally biased trials are sufficient to guarantee an absolute estimation accuracy not exceeding 10% with 95% confidence. For 2-d biasing, the simulation gain will be somewhat lower but the essential advantages of the method above remain. That is, handling a few hundred inversions (once) is not at all a tall order. This method can produce such an avalanche of results that it is tempting to think of it (with a slight stretch of imagination) as a 'turbo-IS'. The above ideas certainly need quantification but it is beyond the remit of this short paper to delve deeper.

An interesting observation comes from Fig. 4, which shows that the biasing parameter is very close to unity and has a small spread despite the range of FAPs considered. The implication is that the (one-sided) density of the metric D has small variance, presumably owing to the choice of L and N. Smaller values of these constants would probably lead to

larger spread of biasing parameter. In the actual adaptations, a small value of the step-size parameter was used to ensure gradual but safe convergence. This explains their apparently slow nature as seen in the figures. While configuring results for a suite of system parameters, only the first adaptation need be somewhat long; subsequent adaptations can be much shorter as they pick up starting or initial values from the previous one.

5. *CONCLUSIONS*

In this modest work small inroads into the use of adaptive IS algorithms to characterize a STAP detector have been made. The AMF was used as example and results have been quite pleasing. Chief reasons for this are that it was possible to invoke the g-method and inverse IS, find a suitable biasing strategy that could easily be optimized adaptively, and find a way around the difficult task of inverting large matrices several times (as described above). The hope is that applications to other STAP configurations, such as normalized AMF and those that handle non-homogeneous clutter, will also meet with success. But this remains to be seen as we certainly are not in position to predict what subtleties these detection algorithms might throw up. It is clear that IS is still in its infancy, especially insofar as use for characterizing modern detection algorithms is concerned. Simulation experiments recounted here have opened possibilities (at least in the authors' minds!) that need further scrutiny.

APPENDIX

ADAPTIVE ALGORITHMS FOR 2-D BIASING

The I-function is estimated as

$$\widehat{I}(\boldsymbol{\nu}) = \frac{1}{K}\sum_{1}^{K} 1(\mathcal{A})\, W^2(\mathbf{X}, \mathbf{X}_L; \boldsymbol{\nu}); \quad \sim f_\star \qquad (14)$$

and its minimization can be carried out using the 2-dimensional adaptive algorithm

$$\boldsymbol{\nu}_{m+1} = \boldsymbol{\nu}_m - \delta \widehat{\mathbf{J}}_m^{-1} \widehat{\nabla} I(\boldsymbol{\nu}_m) \qquad (15)$$

Here, δ is a step-size parameter used to control convergence, and m is the index of recursion. This is a stochastic Newton recursion. It achieves minimization of $\widehat{I}$ by estimating a solution of

$$\widehat{\nabla} I(\boldsymbol{\nu}) \equiv (\widehat{I}_a \quad \widehat{I}_\theta)^T = \mathbf{0}$$

where $I_a \triangleq \partial I(\boldsymbol{\nu})/\partial a$ and $I_\theta \triangleq \partial I(\boldsymbol{\nu})/\partial \theta$. The estimate of the Jacobian $\mathbf{J}$ (which is the Hessian matrix of I) is given by

$$\widehat{\mathbf{J}} = \begin{pmatrix} \widehat{I}_{aa} & \widehat{I}_{a\theta} \\ \widehat{I}_{a\theta} & \widehat{I}_{\theta\theta} \end{pmatrix}$$

where $I_{xy} \equiv \partial I_x / \partial y$. It is straightforward to show that the various I-functions defined above can be obtained by the notational equations

$$I_x = E_\star\{1(\mathcal{A}) W W_x\}$$

$$\begin{aligned} I_{xx} &= E_\star\{1(\mathcal{A}) W W_{xx}\} \\ I_{xy} &= E_\star\{1(\mathcal{A}) W W_{xy}\} \end{aligned}$$

with various derivatives of the weighting function calculated as

$$\begin{aligned} W_a &\equiv \frac{\partial W}{\partial a} \\ &= \left(N - \frac{A}{a}\right)\frac{W}{a} \\ W_\theta &\equiv \frac{\partial W}{\partial \theta} \\ &= \left(LN - \frac{B}{\theta}\right)\frac{W}{\theta} \\ W_{aa} &\equiv \frac{\partial^2 W}{\partial a^2} \\ &= \left[\left(N - \frac{2A}{a}\right)(N-1) + \frac{A^2}{a^2}\right]\frac{W}{a^2} \\ W_{\theta\theta} &\equiv \frac{\partial^2 W}{\partial \theta^2} \\ &= \left[\left(LN - \frac{2B}{\theta}\right)(LN-1) + \frac{B^2}{\theta^2}\right]\frac{W}{\theta^2} \\ W_{a\theta} &\equiv \frac{\partial^2 W}{\partial a\partial \theta} \\ &= \left(LN - \frac{B}{\theta}\right)\left(N - \frac{A}{a}\right)\frac{W}{a\theta} \end{aligned}$$

and they can be estimated as in (14).

REFERENCES

1. F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerospace and Electronic Systems*, AES-28, NO. 1, January 1992, pp. 208-216.
2. M. Rangaswamy, J. H. Michels, and B. Himed, "Statistical analysis of the non-homogeneity detector for STAP applications," *Digital Signal Processing*, vol. 14, 2004, pp. 253-267.
3. R. Srinivasan, "Some results in importance sampling and an application to detection," *Signal Processing*, vol 65, Issue 1, February 1998, pp 73-88.
4. R. Srinivasan, "Simulation of CFAR detection algorithms for arbitrary clutter distributions," *Proc. IEE, Radar, Sonar and Navigation*, vol. 147, February 2000, pp. 31-40.
5. R. Srinivasan, *Importance Sampling - Applications in Communications and Detection.* Berlin: Springer-Verlag, 2002.
6. J. A. Bucklew, *Introduction to rare event simulation.* New York: Springer, 2004.
7. K. R. Gerlach, "New Results in Importance Sampling," *IEEE Trans. Aerospace and Electronic Systems*, vol. 35, No. 3, July 1999, pp 917-925.
8. D. L. Stadelman, D. D. Weiner, and A. D. Keckler, "Efficient determination of thresholds via importance sampling for Monte Carlo evaluation of radar performance in non-Gaussian clutter" *Proceedings of the IEEE Radar Conference*, Long Beach, CA, April 2002, pp. 272-277.