Adaptive Optimization of a Parametric Receiver for Fast Frequency Hopping

H. E. Tayong, A. M. Beasley, and A. Cole-Rhodes
Department of Electrical & Computer Engineering
Morgan State University
Baltimore, MD 21251 USA
e-mail: {tayong, beasley, acrhodes}@eng.morgan.edu

A. B. Cooper III
Department of Electrical & Computer Engineering
The Johns Hopkins University
Baltimore, MD 21218 USA
e-mail: abcooper@jhu.edu

Abstract — **The reduced fuzzy rank order detector (R-FROD) [3] is a robust diversity combiner for fast frequency fopping multiple access. Its performance is determined by the value of an adaptation parameter γ which represents the degree of spread exhibited by the demodulated spread spectrum data. Efficiency (time complexity) of this adaptation is critical for a receiver to be practically useful. A study is presented of three powerful methods for optimizing this parameter. The methods presented are: the golden section search (GSS) algorithm, a straightforward random search (RS) algorithm, and the newer simultaneous perturbation stochastic approximation (SPSA) algorithm.**

I. Introduction

Fast frequency hopping (FFH) waveforms offer powerful resistance to arbitrary noise, multiaccess interference (MAI), and jamming on mobile wireless channels, while maintaining robust performance. The time and frequency diversity of FFH enhance receiver performance provided that careful attention is paid to the design of the diversity combiner. In general, FFH is inherently near/far resistant.

Previously [4] we introduced the fuzzy rank order detector (FROD), a diversity combiner that weights channel data samples according to their distances from the location of the sample distribution, as measured by an *affinity function,* which is a Gaussian shaped function of the difference between a selected data point and a selected order statistic of the data. In [4] we demonstrated the robustness and overall performance of the FROD and introduced the R-FROD a computationally less complex reduced FROD, which suffers little performance degradation relative to the FROD while requiring substantially less effort. Subsequently [3] we showed that the bandwidth efficiency of the R-FROD is competitive with other diversity combiners for FFH with M-ary frequency shift keying (FFH/MFSK). The "inclusiveness" (width) of the Gaussian affinity function is determined by the parameter γ in the same way that the variance of a Gaussian probability density function controls its width.

To date, R-FROD performance has been demonstrated by manually determining values of γ which minimize the error probability, but successful implemenation of the R-FROD in an adaptive communication system requires an efficient, automatic optimization algorithm. Analytic optimization of the objective or loss function (the probability $\Pr(e)$ of error) of the detector is somewhat intractable because it is the ratio of many *sums of sums* of Gaussian functions [4] and because the noise and MAI may be non-Gaussian and, for the mobile channel, non-stationary.

The need for accurate estimation of parameters in the presence of multiple-access interference (MAI) has led to the development of joint multiuser detectors and parameter estimators. In [5], an adaptive receiver is presented using neural network schemes with mathematically tractable nonlinearities. Adaptive receivers presented in [6] have been developed to combat cyclostationary MAI using linear receivers.

This paper focuses on the adaptation algorithms, studying several methods for achieving the optimizing value of γ to minimize the probability of error. Section II presents the system model used as a testbed for the algorithms, which are discussed in Section III. The results of the study are presented and discussed in Section IV and conclusions drawn in Section V.

II. System Model and The Estimator

Consider a conventional[1] MFSK/FFH system with K users, Q hop frequencies, M modulation tones ($Q >> M$), spreading sequence length L, and hop duration T_h. The output of the frequency dehopper at the receiver is represented by an $M \times L$ *detection matrix* whose values are the non-coherently detected energies in each M-ary frequency bin at each time interval T_h. The K signals are assumed hop synchronous, and interference from $K-1$ interferers appears randomly in the entries of the detection matrix. The channel exhibits flat, Rayleigh fading. The channel model and the $M-$ary hypothesis testing

[1]Paper was prepared through collaborative participation in the Communications and Networks Consortium sponsored by the U.S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon.

[1]In conventional FFH/MFSK, $Q >> M$; in unconventional MFSK/FFH $Q = M$ and the modulation tone hops about among the M allowed values. See, *e.g.,* [10].

problem to which it gives rise are discussed in [4]. The detector is the R-FROD.

The R-FROD accepts as inputs the M elements $\mathbf{x} = (x_1, x_2, \ldots, x_M)$ of the detection matrix and the associated vector $\mathbf{x_L} = (x_{(1)}, x_{(2)}, \ldots x_{(M)})$ of order statistics of the same data. Values of the affinity function $\mu(x_i, x_{(j)}; \gamma) = \exp -\{(x_i - x_{(j)})^2/2\gamma\}$ (Figure 1) populate a square, time-rank matrix $\mathbf{T}$. Each row of $\mathbf{T}$ is

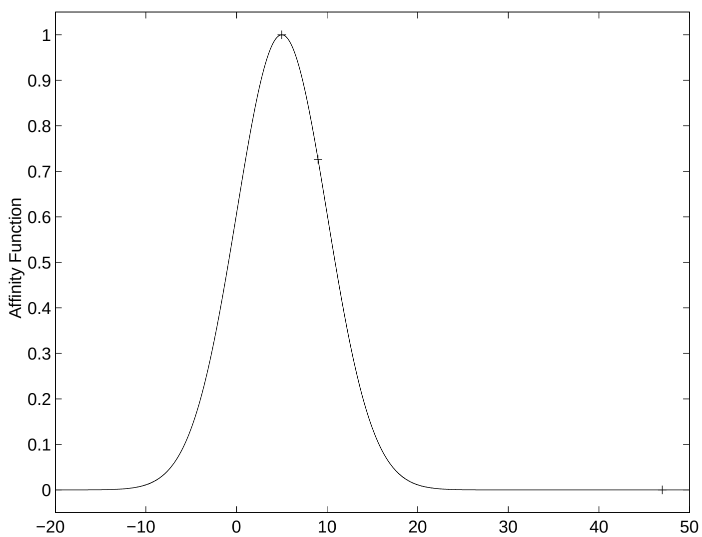

Figure 1: *The Gaussian Affinity Function*

normalized by the sum of its elements to produce matrix $\tilde{\mathbf{T}}$. The *fuzzy ranks* of the data are computed as

$$\begin{aligned} \tilde{\mathbf{r}} &= \tilde{\mathbf{T}}[1, 2, \ldots, M]^T \\ &= (\tilde{\mathbf{r}}_\mathbf{1}, \tilde{\mathbf{r}}_\mathbf{2}, \ldots, \tilde{\mathbf{r}}_\mathbf{M}). \end{aligned}$$

A vector of fuzzy ranks is computed for each column ($M-$tuple) of the detection matrix. Fuzzy ranks extend the concept of (integer-valued) so-called "crisp" ranks for diversity combining [11] by reflecting the statistical spread of the detected energy values. They replace the channel data in the detection matrix. Summation across each row effectively accomplishes the hypothesis testing. More detail is available in [4] and references therein.

Note that a small value of γ causes the detector to discriminate against even close-in "outliers" while a larger value, emblematic of a broader affinity function, gives weight to data points across a broader statistical spread. The objective of this work is to examine algorithms for quickly determining the optimizing value of γ, that value which gives the minimum value of $\Pr(e)$ for the channel environment. The rapid decay of the Gaussian shaped affinity function is intended to discriminate against strong outliers, typical of interference from other nodes in a multiple access system.

In previous work, trial and error were used to determine the optimizing γ and produce the system performance data. Implementation in an adaptive detector, however, will require optimization against a relatively small training set of known transmitted symbols, since the MAI is assumed to be unknown. Thus, in what follows we evaluate candidate optimization algorithms for their efficiency (time complexity) in an adaptive system.

III. Parametric Optimization Algorithms

The adaptive optimization problem is to efficiently determine the value of γ that minimizes the bit error rate using short training sequences of known data. Since the objective function cannot be described by an analytic expression [2], we seek to apply numerical optimization techniques which rely solely on measurements of the objective function and which also do not require a direct measurement of the gradient of the function. We consider recursive optimization techniques which update the solution at each iteration, based upon results of a Monte Carlo simulation of our detector. For a fixed value of the R-FROD parameter γ, the objective function $\Pr(e)$ is determined by averaging over five simulation runs for each case in order to smooth the effects of modeled noise. The shape of the typical plot of error probability vs γ is shown in Figure 2. The presence of a single global minimum makes it feasible to use a simple, one-dimensional search technique such as the golden section search (see A below).

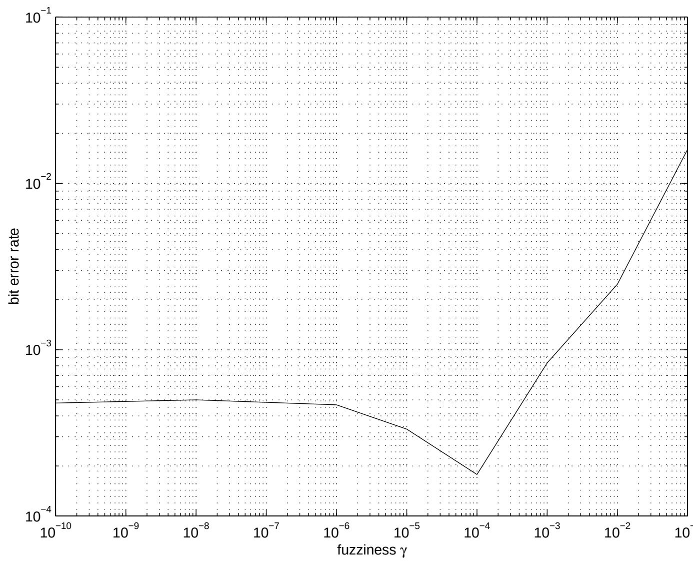

Figure 2: *A typical curve of BER versus γ for the FROD.*

Two other recursive algorithms, which are stochastic in nature, were also applied to this problem. They are the simultaneous perturbation stochastic approximation (SPSA) algorithm and a general random search (RS) algorithm. These are described in the sequel.

A Golden Section Search

GSS [9] is a classical technique which provides a simple method of finding and narrowing the region in which the minimum occurs, in order to achieve convergence to its actual value. Initially, a step size is selected and an interval $[a, b]$ is chosen within which the minimum value of the objective function is expected to lie[2]. To narrow the interval, a point $\bar{x}$ is chosen such that: (a) $a < \bar{x} < b$ and (b) $\bar{x}$ lies a fraction 0.38197 from one end of the interval. This is the so-called golden section ratio [9]. Then $f(\bar{x})$ is evaluated, and the appropriate end of the interval is moved to $\bar{x}$. This algorithm is iterated until the bracket is acceptably small. The nature of this search process makes it optimal for problems having only one true minimum.

B *Random Search*

A random search algorithm is ideal for the present problem, since such an algorithm requires only minimal information about the loss function. The parameter γ is updated at each iteration by $\gamma_{k+1} = \gamma_k + d_k$ where the direction vector d_k is random and is obtained from a normal distribution, $N(0, \sigma^2)$. (In the present case, d_k is a scalar, and a one-dimensional search is, in fact, conducted.) The variable σ must be chosen for best performance, in terms of convergence rate. The value of γ_{k+1} is updated only if it causes the loss function (probability of error) to decrease; otherwise set $\gamma_{k+1} = \gamma_k$. The initial value γ_0 may be chosen from prior or side information or merely selected at random. Random search algorithms thus require function evaluations only at a new value of γ_k and do not require that the loss function be differentiable or that the gradient be computable. They are also useful for loss functions with only one local/global minimum, and these algorithms do not have a lot of random variables that have to be optimized. They are thus easy to implement, and a theoretical proof of performance and accuracy that has been provided for these algorithms in Spall [8].

C *Simultaneous Perturbation Stochastic Approximation (SPSA)*

The simultaneous perturbation stochastic approximation (SPSA) algorithm has recently attracted considerable attention for solving challenging optimization problems where it is difficult or impossible to directly obtain a gradient of the loss or objective function with respect to the parameters being optimized [7]. SPSA is based on an easily implemented and highly efficient gradient approximation that relies on measurements of the objective function, not on measurements of the gradient of the objective function. The gradient approximation is based on only two function measurements (regardless of the dimension of the gradient vector). The update law for the parameter is similar to that of the steepest descent algorithm:

$$\gamma_{k+1} = \gamma_k - a_k g_k.$$

[2]This initial interval is determined from information developed outside the scope of the present problem. If necessary, one need merely think of a very large interval.

where a_k is a sequence of positive numbers that converges to zero and g_k is an approximation to the gradient of $\Pr(e)$. To approximate g_k, we obtain two evaluations of the probability of error:

$$g_k = \{\Pr(\gamma_k + c_k\Delta_k) - \Pr((\gamma_k - c_k\Delta_k)\}/2c_k\gamma_k.$$

where c_k is a decreasing sequence of positive numbers (usually of the form $c_k = c/k^\gamma$, $c > 0$, $0 \leq \gamma < 1$ and where Δ_k takes a value of +1 or 1, as determined by sampling a Bernouilli distribution. See [7] for additional detail. As before, the update law may be initialized by a random "guess" or by prior or side information which brings the algorithm into the neighborhood of the optimum point.

IV. Simulations Results and Discussion

This subsection presents simulation results for the three algorithms. The simulations were conducted under identical conditions. The parameters used to test the algorithms were number of frequencies $Q = 1024$; the number of modulation tones $M = 32$; the number of hops per information symbol $L = 5$. Cases were studied for 64, 73 and 81 users in the system, and the signal to noise ratio (SNR) was 25 dB. The results for the three algorithms are shown in Table 1 below.

Algor.	**# Users**	**# Iter**	**PE**	$log_{10}\gamma$
SPSA	64	9	0.0008	-3.3776
RS	64	29	0.0008	-4.2045
GSS	64	39	0.0007	-4.1530
SPSA	73	7	0.0013	-3.7831
RS	73	45	0.0020	-4.0210
GSS	73	39	0.0021	-4.0167
SPSA	81	4	0.0045	-3.4571
RS	81	27	0.0043	-3.582
GSS	81	16	0.0047	-3.743

Table 1

For each set of users, for the system under consideration, Table 1 shows the number of iterations required by each algorithm to find that value of γ which achieves a specified value of probability of error (PE). The random search (RS) and Golden Section Search (GSS) algorithms each have one variable, equivalent to step-size, which had to be optimized offline, while the SPSA algorithm required optimization of two variables, a_k and c_k. These variables affect performance of each algorithm in terms of convergence rate, and also the accuracy with which it is able to determine the optimizing value of γ at the specified value of error probability $\Pr(e)$.

Each of these three algorithms was fairly simple to implement involving straightforward evaluations of the $\Pr(e)$ by the R-FROD estimator. This was the most computationally intensive part of the optimization algorithms. Thus one consideration in comparing the algorithms is the number calls to the R-FROD estimator, which was required to determine $\Pr(e)$. The random search (RS) and Golden Section Search (GSS) algorithms each required only one function evaluation of $\Pr(e)$ per iteration, while the SPSA algorithm required two function evaluations per iteration in order to determine the gradient. Even so, the SPSA algorithm achieved the optimizing value of γ much faster than the other two algorithms. This is seen in Figure 3, which shows a typical optimization path of γ for each of the three algorithms.

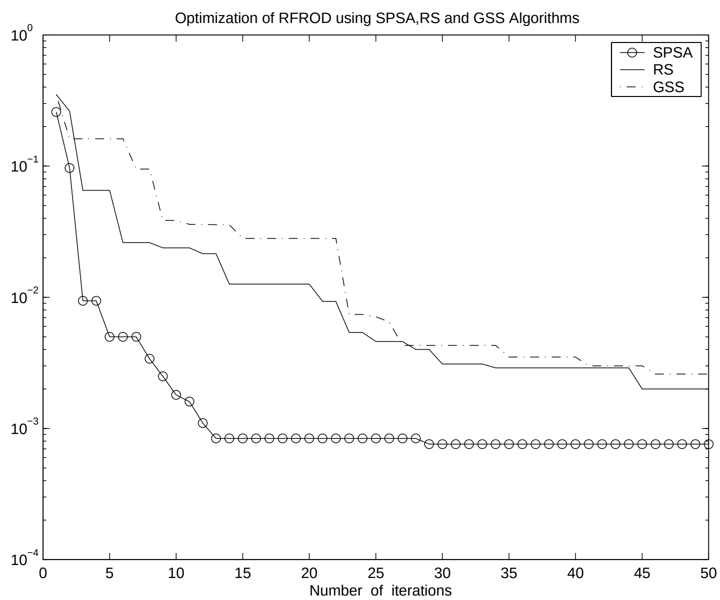

Figure 3: *A typical optimization path of γ for each of the three algorithms*

V. Conclusions

In this work, three stochastic optimization algorithms were investigated and tested on three FFH/MFSK system configurations. The optimization algorithms were iterated to nearly identical values of the objective function, and the value of the system parameter γ was noted. Within the dataset for each system configuration, the final values of error probability never differed by more than 0.0008 and were usually much closer to one another. Values of the logarithm of the optimization parameter were similarly close to one another. (Ideally, they would be identical.)

Within the dataset for each system configuration , the number of iterations required to reach the optimum state was fewest for SPSA by a wide margin. RS took the largest number of algorithmic iterations to complete in two of three cases and GSS the largest in one.

Yet, close examination of the algorithms reveals that RS requires only a single function evaluation in response to random perturbations of the current value of γ_k while GSS requires additional initialization steps plus a more involved updating of the golden section rule. Finally, each iteration of SPSA requires two function evaluations to determine the gradient approximation which is used to update γ_k. Yet, SPSA is the most efficient and finds the optimizing value of γ with the smallest time complexity.

Further and more detailed study is necessary in order to determine the effect on communication efficiency of an adaptive FFH/MFSK multiaccess system using a fuzzy rank order diversity combiner and the SPSA optimization algorithm. Training set size vs performance for various channels must be studied in terms of its impact on bandwidth efficiency and throughput. Meanwhile, we have identified an adaptive algorithm that will improve the efficiency of our ongoing FFH/MFSK research.

Acknowledgments

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government.

References

[1] A. Flaig, G.R. Arce, and K.E. Barner, “Affine order statistic filters: a data-adaptive filtering framework for nonstationary signals,” *IEEE Trans. Sig. Proc.,* vol. 46, pp. 2101-2112, Aug.1998.

[2] A. Flaig, A.B. Cooper III, G. R. Arce, H. Tayong, and A. Cole-Rhodes, “Fuzzy rank-order detectors for frequency hopping networks,” *Proceedings, 3rd Annual Fedlab Symposium on Advanced Telecommunications/Information Distribution Research Program (ATIRP),* pp. 121-125, Feb.1999, College Park.

[3] H. Tayong, A. Beasley, A. Cole-Rhodes, A. B. Cooper, G. Arce, “Bandwidth Efficiency in Fast Frequency Hopping Multiple Access”, *Proc. 35th Annual Conference on Information Sciences and Systems (CISS'01)* Baltimore, MD, March 2001, pp.764-767.

[4] H. Tayong, A. Cole-Rhodes, A. B. Cooper, A. Flaig, G. R. Arce, “Exploiting the Energy Distribution in Non-coherent Frequency-Hopping Diversity Combining”, *Proc. 34th Annual Conference on Information Sciences and Systems,* Princeton, NJ, March 2000.

[5] U. Mitra, H. V. Poor, “Adaptive Receiver Algorithms for Near-Far Resistant CDMA,” *IEEE Trans. Commun.,* vol. 43, pp. 1713-1724, Feb-March-April 1995.

[6] P. B. Rapajic, B. S. Vucetic, “Adaptive Receiver Structures for Asynchronous CDMA Systems,” *IEEE Trans. Commun.,* vol. 12,pp. 685-697, May 1994.

[7] J.C. Spall, “Multivariate Stochastic Approximation Using a Simultaneous Pertubation Gradient Approximation,” *IEEE Trans. Syst, Autom. Control*, 37, 332-341 (1992).

[8] J.C. Spall, J.C, *Introduction to Stochastic Search and Optimization,* (Draft 2001).

[9] W.H. Press, S. Teukolsky, W. Vetterling, B. Flannery, *Numerical Recipes in C,* Cambridge University Press, 1988, available online at http://lib-www.lanl.gov/numerical/index.html.

[10] D.J. Goodman, P.S. Henry, and V. Prabhu, “Frequency hopped multilevel FKS for mobile radio,” *Bell Syst. Tech. J.,* vol. 59, n0. 10, 1980.

[11] T.A. Gulliver, “Order statistics diversity combining in worst case noise and multitone jamming,” in *Conf. Record, 1991 IEEE Military Communications Conference (MILCOM),* (McLean, VA), Mar. 1991.