# BLINDSFTF- STABLE FAST TRANSVERSAL FILTERS ALGORITHM APPLIED TO BLIND EQUALIZATION

*Russell H. Lambert*

University of Southern California
Electrical Engineering-Systems
3740 McClintock Ave. Room 400
Los Angeles, CA 90089
email: rlambert@sipi.usc.edu

## ABSTRACT

SFTF (Stable Fast Transversal Filters), an exact least squares adaptation algorithm with complexity of order 8 times the number of filter coefficients, has been applied to the solution of blind equalization cost functions with dramatic improvements in convergence speed compared to stochastic gradient (lms style) adaptation. This method requires only four times more computation than stochastic gradient methods and has the advantage of being eigenvalue spread independent. The SFTF exact least squares adaptation removes the eigenvalue spread of the channel problem and leaves us now with the speed of convergence depending only on the type of signal used and the inital ISI caused by the channel. Gray's optimum cost function for the generalized Gaussian pdf is discussed and used in the simulations. In summary, an optimal blind cost function is combined with an optimal filter adaptation method. The topic of entropy is discussed and it's relation to kurtosis and speed of convergence properties.

## 1. INTRODUCTION

Using the typical blind equalization problem shown in figure 1, our goal is to adapt the equalization filter $h^{-1}$ so that the overall system transfer function will be a delta function.

The commonly used adaptation method is the stochastic gradient or lms type. Several researchers have proposed that an exact least squares update be used for this blind equalization problem. Before Slock and Kailath introduced the SFTF algorithm in [1], several other researchers had shown that exact least squares adaptation provided both eigenvalue spread independence and increased convergence speed in the blind equalization problem over the simple lms type adaptation methods. Agee [4], Gooch [5], and Weerackody and Kassaam [6], tested blind equalization cost functions with fast recursive least squares (RLS) and lattice-based adaptation algorithms.

This work was supported in part by the Office of Naval Research under contract No. N00014-91-J-4021

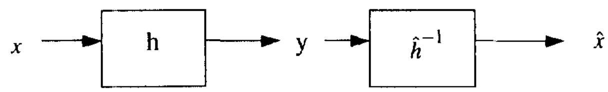


Figure 1: Block Diagram of Blind Equalization System

The SFTF algorithm has the smallest computational requirement of those studied previously and yet retains speed of convergence as well as stability. Papadias and Slock in [2] have presented the feasibility and benefits of the SFTF approach to blind equalization using constant modulus cost functions. This paper presents the benefits of the SFTF algorithm adaptation using Gray's Variable Norm [7], which is applicable to many data types, as the cost function.

## 2. MOTIVATION: NEWTON-TYPE UPDATE FOR BLIND EQUALIZITON COST FUNCTIONS

Denoting the blind cost function used as $J_b$, through calculation of the expectation of the second derivative of the cost function $\frac{\partial^2 J_b}{\partial^2 h}$, one can obtain a Newton-type, quickly converging update procedure. In the case of adaptive filtering using the reference signal, one obtains $R$ the correlation matrix to be the expectation of the second derivative and the update is called the exact least squares approach.

In the case of blind equalization, one obtains a similar expression which also contains the correlation matrix $R$. This expression can be approximated by using $\mu^{-1}R$, where $\mu^{-1}$ is simply a constant. This result



1058-6393/95 $4.00 

means that the use of an adaptation algorithm which was derived for the known-reference problem could be used in the blind equalization problem with only minor modification.

The eigenvalue spread independence of the new technique is demonstrated in the simulations. The speed of convergence as compared the stochastic gradient methods is quite dramatic. A comparison of the performance of the SFTF approach, of order $8N$ complexity, to the "lmsstyle" steepest descent or stochastic gradient approach of order $2N$ complexity, can be made by refering to [12]. Some background on the cost functions used in this study is given. Then a description of the SFTF algorithm for use with blind cost functions will be presented, followed by simulation results and conclusions.

## 3. BLIND EQUALIZATION COST FUNCTIONS AND ENTROPY CONSIDERATIONS

Any signal, which is i.i.d. and has finite variance, will become "Gaussianized" due to the Central Limit Theorem when it passes through a filter. Thus blind equalization cost functions assume that the pdf of the original source signal has been Gaussianized by the channel and attempt to measure the Gaussianity of $y$ and drive the signal as far as possible from Gaussian. One scale invarient measure of Gaussianity is normalized kurtosis:

$$\frac{Ex^4}{(Ex^2)^2}, \tag{1}$$

which can be generalized to form what is known as Gray's Variable Norm [7],in order to accommodate more optimum norms for different types of input signals:

$$O_s^r = \frac{E|\hat{x}|^r}{(E|\hat{x}|^s)^{\frac{r}{s}}}. \tag{2}$$

Gray's cost function can also be thought of as a measure of entropy, Walden [8], discusses it's relation to entropy. Gray's cost function $O_\alpha^2$, has been shown to be optimum for the equalization of a signal which can be modeled with the generalized Gaussian pdf [7], given by

$$P_x(x) = \frac{\alpha}{2\beta\Gamma(\frac{1}{\alpha})}e^{-(\frac{|x|}{\beta})^\alpha}, \tag{3}$$

where $\frac{\alpha}{2\beta\Gamma(\frac{1}{\alpha})}$ is just the normalization constant and $\alpha$ is the shape parameter. Using a generalized Gaussian pdf as the input source, Gray's function $O_\alpha^2$, is the optimum blind equalization cost function from the Hypothesis Testing [7] or Maximum A Posteriori estimation point of view [11]. In fact, Shalvi-Weinstein in [15], have also derived the gradient of Gray's cost funtion for two specific values of $s$ from a completly new approach. Gray's variable norm $O_\alpha^2$ has therefore, been in some sense, derived in three ways and is shown to be optimal for the generalized Gaussian pdf by two of those methods.

The generalized Gaussian is useful in parameterizing many pdfs. Special cases are: $\alpha = 1$, the Laplace (or double-sided exponential), $\alpha = 2$, the gaussian, and $\alpha = \infty$, the uniform distribution.

## 4. BLINDSFTF: SFTF ADAPTATION OF A BLIND COST FUNCTION

Details of the FTF(fast transversal filters) approach are described in [8]. FTF calls for adapting four filters instead of one, a forward predictior, a backward predictor, a kalman gain vector, and the true inverse filter itself. Slock and Kailath presented an excellent method of introducing redundancy and stablizing the FTF, calling it SFTF. Essentially the SFTF algorithm given by Slock and Kailath can be used for the blind equalization problem, but instead of using

$$J_{nb} = (x - \hat{x})^2, \tag{4}$$

as the "non-blind" cost function, which gives a gradient of $(x - \hat{x})y$, we used a blind cost function

$$J_b = O_s^r = \frac{E|\hat{x}|^r}{(E|\hat{x}|^s)^{\frac{r}{s}}}. \tag{5}$$

The gradient for $O_s^r$ is

$$\frac{\partial O_s^r}{\partial \hat{h}^{-1}} = sign(\hat{x})\ (|\hat{x}|^{r-1} - \frac{E|\hat{x}|^r}{E|\hat{x}|^s}|\hat{x}|^{s-1})y. \tag{6}$$

Due to central limit theorem arguments stated above, we usually chose $r = 2$ and the gradient will take the simpler form:

$$(\hat{x} - g(\hat{x}))y. \tag{7}$$

## 5. NEWTON UPDATE FOR BLINDSFTF WITH $O_4^2$

In BLINDSFTF, an additional step size is needed to scale down the error from the blind cost function. This additional step size can be found by using Newton's update which uses the expectation of the inverse second derivative of the cost function $(E\{\frac{\partial^2 J}{\partial^2 \hat{h}^{-1}}\})^{-1}$.

The Newton-type of update for $J_{nb} = (x - \hat{x})^2$, (the non-blind known reference case) is:

$$\hat{h}^{-1} = h^{-1} + R^{-1}\frac{\partial J_{nb}}{\partial \hat{h}^{-1}}, \tag{8}$$

A Newton-type of update for $J_b = O_4^2$, is:

$$\hat{h}^{-1} = \hat{h}^{-1} + \left(\frac{1}{1 - \frac{3Ex^2\ E\hat{x}^2}{Ex^4}}\right) R^{-1} \frac{\partial O_4^2}{\partial \hat{h}^{-1}} \qquad (9)$$

If a simple constant step size is to be used instead of $\left(\frac{1}{1-\frac{3Ex^2\ E\hat{x}^2}{Ex^4}}\right)$, then a useful upper bound for the additional step size $\mu$ would be:

$$\mu < \frac{1}{1 - \frac{3Ex^2\ E\hat{x}^2}{Ex^4}}. \qquad (10)$$

Similar types of bounds can be found for other specific cases. Note that in (10) the term $E\hat{x}^2$ is a measure of channel dispersion, i.e. the varience of the input to the equalizer. It is also the initial value of ISI + 1. Also of interest is the ratio $\frac{Ex^2\ E\hat{x}^2}{Ex^4}$, which contains a measure of entropy by comparing it with (1). Thus we see that a "mean error weight vector analysis" for a *blind* adaptation would contain three factors,

1. Eigenvalue spread term, determined by the eigenvalues of the $R$ matrix.
2. Entropy of the input source.
3. Channel dispersion or initial level of ISI.

The upper bound of $\mu$, which is the maximum step size the algorithm will allow without going unstable, is a rough measure of the power of the cost function with respect to the baseline known-reference adaptation. The SFTF algorithm has excellent stability and speed. The known reference version can usually be tuned so that convergence is attained in a few filter lengths. In the simulation carried out in this study, the additional step size seemed to be a good measure of performance compared to the adaptation using a reference. As an example of this, when the upper bound on $\mu$ was found to be .25, the convergence seemed to take four times longer.

Interestingly, there are certain cases using very low dispersion channels and low entropy signals (such as 2-PAM) where the upper bound is greater than one, implying that the blind cost function $J_b$ is more powerful than $J_{nb}$, the known reference cost.

## 6. SIMULATIONS

Two linear channel models were used in the testing of BLINDSFTF. Channel 1 was $[1\ 1\ -.75]$ with eigenvalue spread of 7 and channel 2 was $[.4\ 1\ -.7\ .6\ .3\ -.4\ .1]$ with eigenvalue spread of 28. Three types of data were used, Laplace using $O_1^2$, 2-PAM data using $O_4^2$ and uniformly distributed data also using $O_4^2$. For all the tests, the time constants and step sizes where chosen so that the steady state level of ISI was .015. The $\lambda$ parameter was always chosen as $\lambda = 1 - 1./t$, where $t = 4M$ and M is the filter order. The AMU constant which controls the initial diagonal on $R^{-1}$ was 700.

| Distribution | Alpha | Kurtosis $\frac{E\lvert x\rvert^4}{(E\lvert x\rvert^2)^2}$ | Entropy with $\sigma^2 = 1$ |
|---|---|---|---|
| 2-PAM | NA | 1 | .69 nats |
| Uniform | $\infty$ | 1.8 | 1.24 nats |
| Laplace | 1 | 6 | 1.346 nats |
| Gaussian | 2 | 3 | 1.4189 nats |

Table 1: Signal Types Used in Simulations Compared to Gaussian

BLINDSFTF can be compared to the traditional lms type of blind equalizer adaptation by refering to [12]. In these tests BLINDSFTF converged from 3 to 20 times faster. Laplace data was the most difficult, and test results of channel 2 using Laplace data which converged in about 5000 iterations using BLINDSFTF (plotted in Fig. 8), required over 80000 iterations using the lms type update. These tests have shown a convergence speed vs. entropy characteristic shown in figure 2. Theorectically, as the entropy of the source data approaches 1.4189nats (that of the Gaussian), the number of iterations needed for convergence is infinity. This fact is also predicted by the bound given in (10)

$$\frac{1}{1 - \frac{3Ex^2\ E\hat{x}^2}{Ex^4}},$$

since the kurtosis of the Gaussian is 3.

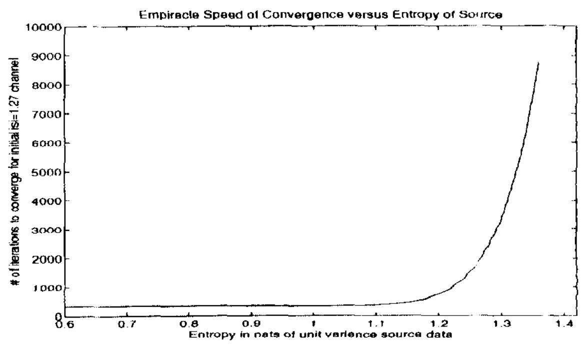


Figure 2: Empiracle convergence speed vs. entropy of source data using channels of initial ISI=1.27 and BLINDSFTF which removes the eigenvalue spread sensitivity.

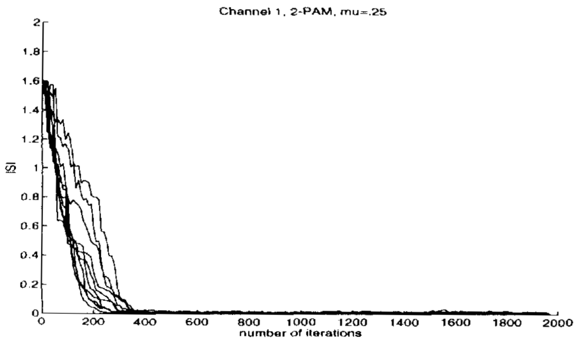


Figure 3: Convergence of ten realizations of 2-PAM data using BLINDSFTF and channel 1 with eigenvalue spread 7.

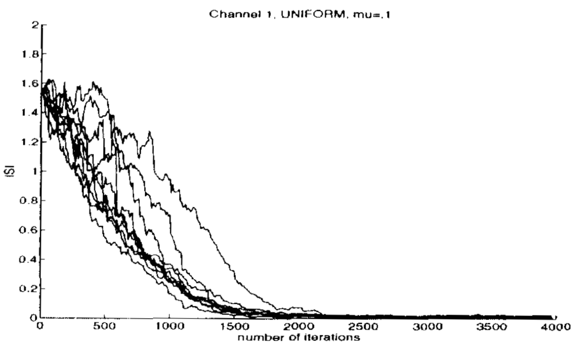


Figure 4: Convergence of ten realizations of UNIFORM data using BLINDSFTF and channel 1 with eigenvalue spread 7.

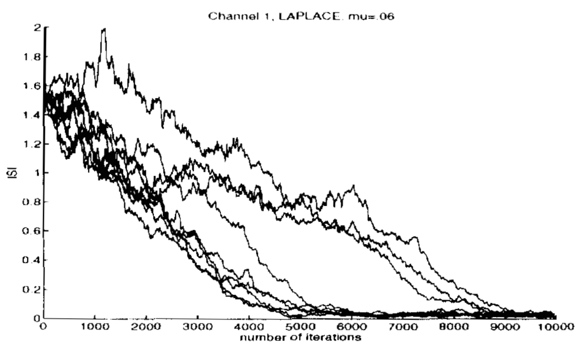


Figure 5: Convergence of ten realizations of LAPLACE data using BLINDSFTF and channel 1 with eigenvalue spread 7.

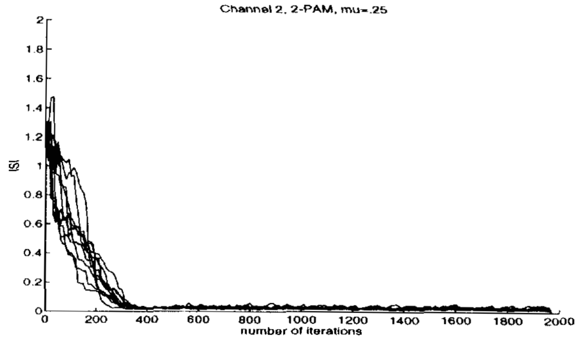


Figure 6: Convergence of ten realizations of 2-PAM data using BLINDSFTF and channel 2 with eigenvalue spread 28.

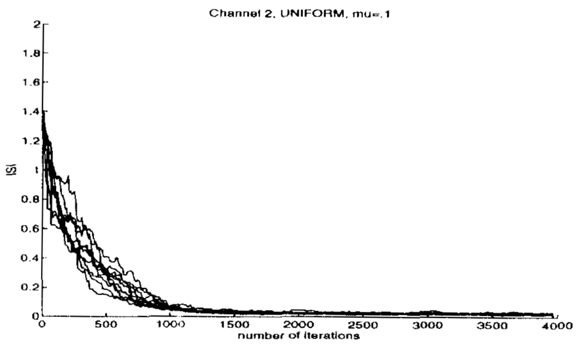


Figure 7: Convergence of ten realizations of UNIFORM data using BLINDSFTF and channel 2 with eigenvalue spread 28.

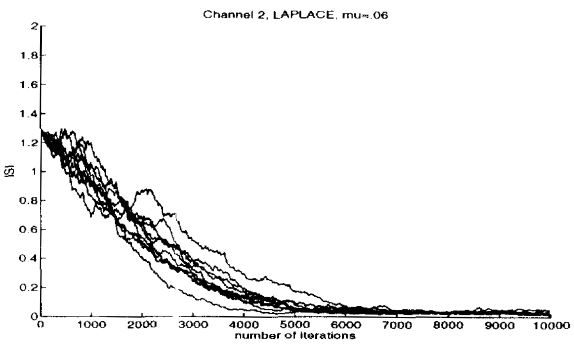


Figure 8: Convergence of ten realizations of LAPLACE data using BLINDSFTF and channel 2 with eigenvalue spread 28.

## 7. CONCLUSION

In blind equalization there are three factors that influence convergence speed: distance from Gaussian (or entropy) of the source data itself, eigenvalue spread of the channel and channel dispersion (or initial ISI). This is in contrast to known reference adaptation where only channel eigenvalue spread is the problem for non-exact least squares adaptation. An powerful new tool for blind adaptive filtering has been tested. SFTF applied to blind cost functions provides optimal exact least squares convergence at the expense of only 4 times lms complexity.

## 8. REFERENCES


[1] Dirk T. M. Slock and T. Kailath, "Numerically stable fast transversal filters for recursive least squares adaptive filtering", IEEE Trans. ASSP Vol. 39, pp. 92-115, Jan. 1991.

[2] Constantinos B. Papadias and Dirk T. M. Slock, "New adaptive blind equalization algorithms for constant modulus signals", ICASSP 94, Australia, pp. III-321-324

[3] O. Shalvi and E. Weinstein, "Super-exponential methods for blind deconvolution", IEEE Info. Theory, Vol. 39, March 1993, pp.504-519.

[4] Brian G. Agee , "The least squares CMA: a new technique for rapid correction of constant modulus signals", ICASSP 86, Tokyo, pp. 293-296

[5] Richard P. Gooch , "A lattice-based constant modulus adaptive filter", Twentieth Asilomar Conference on Circuits, Systems, and Computers, Nov. 86.

[6] Vijitha Weerackody and Saleem A. Kassaam, "Blind equalization using lattice filters", ICASSP 88.

[7] William Gray, "Variable Norm Deconvolution", Ph.D. Dissertation Stanford University, Aug. 1979.

[8] A. T. Walden, "Non-Gaussian reflectivity, entropy and deconvolution", Geophysics, Vol. 50, Dec. 1985, pp. 2862-2888.

[9] Simon Haykin, "Adaptive Filter Theory", McGraw-Hill, second edition.

[10] E. H. Satorius and J. J. Mulligan, "An alternative methodology for blind equalization," *Digital Signal Processing*, Academic Press Review Journal, vol. 3 no. 3, pp. 199-220, 1993.

[11] Russell H. Lambert and Chrysostomos L. Nikias, "Optimal equalization cost functions and maximum a posteriori estimation," MILCOM 1994.

[12] Russell H. Lambert and Chrysostomos L. Nikias, "Forward/inverse blind equalization," ASILOMAR Signals and Systems Conference 1994.

[13] R. V. Hogg,"More light on the kurtosis and related statistics," *Jounal of American Stat. Soc.*, vol. 67 no. 338, pp. 422-425, June 1972.

[14] V. A. Uthoff, "An optimum test property of two well-known statistics," *Jounal of American Stat. Soc.*, vol. 65 no. 332, pp. 1579-1600, Dec 1970.

[15] O. Shalvi and E. Weinstein, "Super-exponential methods for blind deconvolution," *IEEE Trans. on Info. Theory*, vol. 39, pp. 504-519, March 1993.