DTIC FILE COPY

DREP Technical
Memorandum 88-9

(2)

Defence Research Establishment Pacific

Centre de recherches pour la défense pacifique

DEFENSIO PER SCIENTIAM

AD-A196 759

Maximum Entropy Criteria Applied to Signal Recovery

Robert F. MacKinnon and Michael J. Wilmut

June 1988

Research and Development Branch
Department of National Defence

Canada

Defence Research Establishment Pacific
Centre de recherches pour la défense pacifique

DEFENCE RESEARCH ESTABLISHMENT PACIFIC

CFB Esquimalt, FMO Victoria, B.C. VOS 1BO

DREP Technical Memorandum 88-9

MAXIMUM ENTROPY CRITERIA APPLIED TO SIGNAL RECOVERY

Robert F. MacKinnon and Michael J. Wilmut

June 1988

Approved by:

CHIEF

Research and Development Branch
Department of National Defence

Canada

ABSTRACT

A method based on the minimization of cross-entropy is presented for the recovery of signals from noisy data either in the form of time series or images. Finite Fourier transforms are applied to the data and constraints are placed on the magnitude and phase of the Fourier coefficients based on their statistics for noise-only data. The minimization of cross-entropy is achieved through application of well-established functional minimization techniques which allow for further constraints in the spatial, temporal or frequency domain. Derivatives of the entropy function are obtained analytically and the results applied to the cases of correlated noise and of signal perturbations about a mean. Demonstrations of applications to one-dimensional data are presented.

I. INTRODUCTION

A common problem of data analysis is the reconstruction of a signal from noisy data. The objective is to obtain as true a representation as the noisy data will allow without having exact knowledge of the event giving rise to the signal or of its consequences.

Usually some knowledge of the signal is available. For example if the data consist of a radar image of the ocean surface and the signal sought is an internal wave pattern, then a great deal is known about the physical properties of such waves. In some cases independent sea-surface measurements are available which allow estimation of wavelengths, amplitudes and velocities. The reconstruction of the wave pattern will be deemed unsatisfactory unless it conforms to the background knowledge. What is desired in such cases is a method of incorporating prior knowledge in the signal recovery process while maintaining a degree of flexibility consistent with the state of that knowledge and the reliability of the data.

In this paper a method of signal reconstruction is described which evokes the principle of minimum cross-entropy, a generalization of the principle of maximum entropy, and which incorporates prior knowledge in the form of constraints on the solutions to minimization problems. The method uses the finite Fourier transform and constraints may be applied in either the spatial or frequency domain. The solutions described herein were obtained using a general purpose minimization routine and were restricted to one-dimensional data. For two-dimensional data, a faster special-purpose routine has been developed based on the same principles [1]. In discussing the method, reference will be made to image processing since it was this application for which it was envisioned. Analytical results given for the one-dimensional case can be extended easily to two-dimensions.

II. ENTROPY AND THE WIENER FILTER

A digital time series, or image, which is the result of a stochastic process can be considered as a set of values, d(x), of limited extent and accuracy which are assigned to a discrete set of locations x_i, i=0,1,...M-1. The values are required to be non-negative and are taken to be representative of intensities. The individual intensities can be considered to be the proportion of the total available intensity which is assigned to a particular location. The problem at hand is to arrive at an estimation of the signal values, s(x), when the data could be fitted equally well by many estimation sets.

Given that this problem is ill-posed in the Hadamard sense, various schemes are available for defining an associated problem which is well-posed [2]. However, no method can be said to be optimum in the wide sense, so the selection of a particular method depends on the nature of the knowledge available as well as on the experience of the analyst. When noise is known to be a major component of the data, a desirable feature is the availability of a related measure of statistical significance which ideally manifests itself as a parameter defining a family of solutions.

Entropy optimization is a powerful, general technique which provides solutions which possess many desirable properties [3][4]. Maximum entropy implies "maximally smooth" and sometimes "maximally likely" as well. Such solutions are said to be the simplest possible result containing the bare minimum of structure needed to fit the constraints imposed by the data [5]. Furthermore it has been shown that optimum generalized entropy solutions possess uniquely certain properties of consistency in cases where the given data are supplemented subsequently [6]. Although the methods of maximum entropy and its generalization, minimum cross-entropy, may be subject to various interpretations, there is merit in the approach of Frieden [7] based on information theory since this provides a useful context for analyzing attempts at signal reconstruction. A brief summary follows.

The classical (Shannon) definition of the information, $I(A,B)$, contained in a message A about an event B is given by

$$I(A,B) = \ln\left[P(B|A)/P(B)\right] \tag{1}$$

where P represents the probability density function.

The event B may be taken to be a process which produces a continuous variable intensity, s, and the message A may be associated with a set of measurements, d, of the consequences of the event. The entropy, H, of the event is defined as

$$H(B) = -\int_{-\infty}^{\infty} P(y) \ln P(y)dy. \tag{2}$$

The performance of a measuring device is difficult to characterize in general. In practice often it may be considered to be band-limited, linear in response, and noisy. A common model describing its performance is

$$d(x) = s(y) \times g(y) + n(x) \quad , \tag{3a}$$

where d represents the data
s represents the signal
n represents the noise, independent of s,
g represents the action of the instrument
and x represents the process of convolution through which this action is accomplished.

A Fourier transform of this equation yields

$$D(\omega) = S(\omega)\, G(\omega) + N(\omega) \equiv Y(\omega) + N(\omega) \tag{3b}$$

where ω represents frequency, either in a spatial or temporal sense, and the capital letters represent the Fourier transforms of the corresponding lowercase variables.

If restoration is sought through the multiplication of $D(\omega)$ by a weighting function, $W(\omega)$, and a subsequent inversion of the resulting transform data, it is well-known that the minimum expected mean-square error in the restored signal is achieved by the so-called Wiener filter given by

$$W(\omega) = G^{-1}(\omega) \left[r(\omega) / \{1 + r(\omega)\}\right] \tag{4}$$

where $r(\omega) = |Y(\omega)|^2/|N(\omega)|^2$, which represents the data signal-to-noise power ratio at frequency ω.

Frieden [7] demonstrates that this filter may be derived from information-theoretic principles, in particular, through maximization of the transinformation, $I(Y,D)$, defined as the difference between the entropies associated with $Y(\omega)$ and $N(\omega)$. These can be shown to be functions only of the power spectral coefficients given by

$$H(Y) = 1 + \ln(\pi|D|^2) \tag{5}$$

and

$$H(N) = 1 + \ln(\pi|N|^2)$$

under the assumption that the noise is Gaussian of zero mean and independent of the signal.

The maximum possible $I(Y,D)$ is termed the channel capacity, $C(\omega)$, and is a measure of the potential for restoration of the signal component at frequency ω under the given noise conditions. It can be shown that

$$C(\omega) = \ln(1+r(\omega)) \tag{6}$$

so that $W(\omega) = \left[1 - \exp\{-C(\omega)\}\right]/G(\omega)$.

In the present paper the main concern is noise reduction, so $G(\omega)$ may be taken to be unity. In this situation the Wiener filter approach is to apply weights to the Fourier coefficients, thereby reducing those coefficients where the signal-to-noise ratio is expected to be low (low channel capacity) while leaving relatively unaffected those coefficients whose signal-to-noise ratio is expected to be high. Phases are preserved at all frequencies whereas some loss of signal power can be expected on average, the cost of the improvement in overall signal-to-noise ratio.

III. RESTORATION IN THE FOURIER DOMAIN

When a signal has undergone a convolution process it is convenient to treat the data after Fourier transformation. Even if, as herein, the major concern is not deconvolution, there are good reasons for frequency domain processing. When noise-only samples undergo linear transformation, the resultant coefficients are often nearly independent Gaussian distributed as a consequence of the central limit theorem [8]. Therefore, amplitude and phase statistics are well approximated by well-known distributions regardless of the noise distribution in the spatial domain. In other words, frequency domain processing tends to be robust.

Wiener filtering is widely used for restoration [9], even when knowledge of the signal spectrum is inexact. The analyst must accept less than optimal results, which may or may not be satisfactory. Often crude approximations to the signal spectrum are sufficient. For unknown signals some property may be chosen as characteristic of a desirable solution to serve as a basis for a criterion of optimality. Candidate solutions are constrained to conform within limits based on the noise statistics. Methods of least squares constraints seek solutions for which the estimated error variance matches that of the noise.

Hunt [10] developed a method for which the criterion of optimality was the minimization of the square of the Laplacian - a smoothness criterion implying that the signal was basically low frequency in content. Gull and Daniel [11] employed another criterion for smoothness - maximum entropy - combined with a measure of expected error. Their method seeks solutions which minimize $Q(\lambda)$ defined by

$$Q(\lambda) = -\sum_j f_j \ln f_j - \lambda \sum_\omega |Y(\omega) - D(\omega)|^2/\sigma^2(\omega) \qquad (7)$$

where f_j represents intensity of the j-th location

λ represents a positive Lagrangian multiplier

Y represents the estimated transform

D represents the data transform

σ represents the noise standard deviation

and the sum over ω is taken over a set of preselected frequencies.

The solution is obtained by an iterative procedure, the correct value of λ being that for which the sum over ω achieves its expected value. Candidate solutions can be assigned confidence levels based on the noise statistics.

Direct subtraction of noise power in the frequency domain is a simple method of noise reduction, perhaps best described by Lim [12]. The estimated signal Fourier coefficient, $Y(\omega)$, is defined by

$$\begin{aligned} |Y(\omega)|^2 &= |D(\omega)|^2 - \alpha\,\sigma^2(\omega), \text{ if positive} && (8) \\ &= 0 \qquad\qquad\qquad\quad , \text{ otherwise.} \end{aligned}$$

The phase of $Y(\omega)$ is taken to be the phase of $D(\omega)$. If $D(\omega)$ equals $S(\omega)+N(\omega)$, then $|D(\omega)|^2/\sigma^2(\omega)$ is well approximated by a noncentral chi-squared distribution of mean $|S|^2/\sigma^2+2$ and variance $4|S|^2/\sigma^2+4$ [13]. Normalization of Eq. (8) yields

$$|Y(\omega)|^2/\sigma^2(\omega) = |D(\omega)|^2/\sigma^2(\omega) - \alpha \qquad (9)$$

where the left-hand side has a mean-shifted noncentral chi-squared distribution. If $\alpha > 2$, the expected value is less than $|S|^2/\sigma^2$, so some loss of signal power results on average. On the other hand, the larger α, the more likely that the residual power is due to the presence of a signal at the given frequency.

The loss of signal power under these circumstances may not be critical since it is the phase spectrum which determines the general features of an image [14]. Acceptance of the phase spectrum of the data implies acceptance of these general features which, in turn, suggests a high signal-to-noise ratio. The accuracy of the phase estimate is a function of the signal-to-noise power ratio and therefore of the channel capacity at a given frequency. It is shown below that cross-entropy is a function of the phase when prior spatial information on the signal is introduced.

For the method proposed in the next section, constraints are applied in both the spatial and frequency domains while Fourier magnitudes and phases are allowed to vary independently. Obviously constraints may be chosen so that no solution exists. A trivial example would be the case in which only a single frequency component is allowed while the signal is required to be zero over a certain interval. Furthermore, it is well-known that if a signal is known to satisfy certain constraints, that signal may be specified uniquely by partial Fourier domain information such as the transform magnitude [15] or phase or even the sign of its real part [16]. These issues of existence and uniqueness are a subject of continuing research [2].

IV. A CONSTRAINED MINIMUM CROSS-ENTROPY METHOD

In this section the proposed method is described. Assume the data are a set of intensities denoted by $d_j, j=0,1,2,\ldots.M-1$, which are the sums of a signal component, s_j and an independent noise component, n_j. It is desired to find a representation of the signal which is consistent with the data, given estimates of the means and variances of the noise power spectral coefficients and the phase spectrum at frequencies $\omega_j, j=0,1,..(M/2)+1$. (The phase data are expressed best in terms of the arctan of the estimated phase which tends to be normally distributed [8]). The proposed method of solution is as follows. Obtain the finite Fourier transform of the data. $D(\omega_j)$. Let $S(\omega_j)$ represent the estimate of the Fourier coefficient of the signal at ω_j. Select as a first estimate of $|S(\omega_j)|^2$ the value of $|D(\omega_j)|^2 - |N(\omega_j)|^2$, if positive, otherwise assume it to be zero. The first estimate of the phase is the phase of $D(\omega_j)$. Place bounds on the subsequent estimates of magnitude and phase using the prior knowledge of the noise spectral properties. Now seek a solution which minimizes the generalized entropy. This process is either the maximization of entropy or the minimization of cross-entropy, as appropriate. It is assumed that the sum $\sum s_j$ is equal to a given constant and that the signal components are all positive.

To achieve a solution a general purpose quasi-Newton minimization routine was chosen from the NAG library [17]. If the bounds on the frequency domain variables are "tight", this implies that the signal-to-noise ratio is high and the solution is required to conform closely to the data. If the bounds are "loose" the solution may depart significantly from the data values, now assumed to be noisy, and will tend to conform to the prior knowledge.

Additional constraints may be applied on a point-by-point basis in either domain, so prior knowledge can be incorporated into the solution with ease. If, for example, it is known that the signal has a certain minimum total power, the solution may be constrained so that this requirement is fulfilled. If the signal is known to decrease in amplitude as j increases, this may be translated into a decreasing upper bound in the original data domain.

Any initial estimate of $S(\omega_j)$ is acceptable provided only that it lies within the prescribed range. Generally, Wiener filter solutions provide initial estimates closer to the final solution than do those obtained from the noise subtraction process described above. This is to be expected since the Wiener filter is based upon an exact knowledge of the signal spectrum.

It should be noted that constraints in the frequency domain can be applied to the real and imaginary parts of the transform coefficients rather than to their phase and magnitude. Results presented elsewhere [1] were obtained in this manner. The domain of acceptability is not the same for the two methods.

Examples of the application of the algorithm to one-dimensional simulated data are given in Section VIII. The next three sections present analytical results which provide some insight into the process and the solutions expected.

V. ENTROPY DERIVATIVES

The quasi-Newton minimization technique requires knowledge of the derivatives of the objective function with respect to the constraint variables [17]. The derivatives of the entropy function with respect to Fourier magnitude and phase can be obtained in a straightforward manner from the properties of the finite Fourier transform. For a function, f, defined at a discrete set of equispaced values $f_j, j=0,1,2,..M-1$, its transform, $F_m(f)$, may be defined as

$$F_m(f) = \sum_{j=0}^{M-1} f_j \exp[-i(2\pi jm/M)] \tag{10}$$

$$= M[a_m \exp i\phi_m]$$

$$= M[u_m + iv_m]$$

where a_m and ϕ_m are thus defined in terms of the magnitude and phase of the Fourier coefficients. Let the generalized entropy function, H, be defined by

$$H = \sum_{j=0}^{M-1} f_j \ln(f_j/b_j) \tag{11}$$

where the b_j's represent the assumed values of the solution. For uniform b_j, H is the negative of the sample entropy, whereas in general H is the sample cross-entropy. H and f_i may be defined as functions of the a_m and ϕ_m by applying inverse Fourier transformations to Eq. (10).

For m=1,2,...,(M/2)-1, by differentiation of Eq. (11) it can be shown that

$$[\partial H/\partial \ln(a_m),\ \partial H/\partial \phi_m] = 2M^{-1}F_m(w) \tag{12}$$

where $w_k = \sum_{j=0}^{M-1} f_j \ln(f_{j+k}/b_{j+k})$.

and where the quantities inside the square brackets represent the real and imaginary parts of a complex quantity, respectively.

For m=M/2, the factor 2 is replaced by unity. The derivative for m equal to zero is not required since the mean value of the f_j is fixed. Furthermore, from the convolution property of the transform, it follows that

$$F_m(w) = F_m^*(f)F_m(\ln(f/b)) \ , \tag{13}$$

where the asterisk denotes complex conjugation.

Generally, $\partial H/\partial \phi_m$ is non-zero since the phase of $F_m(f)$ will not match exactly that of $F_m(\ln f/b)$, even if b is uniform.

The derivatives with respect to the Fourier coefficients are given by

$$[\partial H/\partial u_m,\ \partial H/\partial v_m] = 2M^{-1}F_m(\ln(f/b)), \tag{14}$$

for m=1,2,..,(M/2)-1. The derivatives become zero when the transform coefficient for ln(f/b) becomes zero, whereas the derivatives with respect to magnitude and phase are zero also when $F_m(f)$ is zero, independent of the prior knowledge. For generalized entropy the solutions sought are those for which the variations in f match those of b.

By definition w_0 is the cross-entropy with prior b. Otherwise w_k may be expressed in the form

$$w_k = \sum_{j=0}^{M-1} \left[f_j \ln(f_j/b_{j+k}) - f_j \ln(f_j/f_{j+k}) \right] \quad . \tag{15}$$

The first term represents the cross-entropy for a shifted prior. The second term is a sample "auto-cross-entropy" in which the shifted prior is replaced by the actual value measured.

VI. ENTROPY FOR SIGNAL AND NOISE PERTURBATIONS

Suppose the data, f_j, lie between zero and unity and may be represented as a mean value, μ, minus small variations, z_j, equal to the sum of a signal s_j and zero-mean noise, n_j, independent of the signal. At the j-th location, to order δ^2,

$$\begin{aligned} \ln(f_j/b_j) &= \ln\left[(\mu - z_j)/b_j\right] \\ &= \ln(\mu) - 2\delta_j - 2\delta_j^2 - \ln(b_j) \end{aligned} \tag{16}$$

$$\text{where } \delta_j \equiv (s_j + n_j)/2\mu \ .$$

It can be shown that, for m=1,2,..(M/2)-1,

$$\langle F_m^*(f) F_m(\ln(f/b)) \rangle = 4\mu \langle |F_m(\delta)|^2 \rangle + 2\mu \langle F_m^*(\delta) \rangle \ F_m(\ln(b)) \ , \tag{17}$$

where < > denotes expected value.

The second term generally has an imaginary part, which introduces a phase dependence to the generalized entropy. However, if the b_j are uniform, then

$$\langle F_m^*(f) F_m(\ln f) \rangle = \mu^{-1}\left[|S_m|^2 + |N_m|^2\right] \equiv M^2 \langle a_m^2 \rangle \tag{18}$$

where S_m and N_m represent $F_m(s)$ and $F_m(n)$, respectively.

To achieve maximum entropy the amplitude at each frequency is reduced, on average, in proportion to the power at that frequency, which is the total of signal plus noise power. This is consistent with Frieden's analysis outlined above since $\ln(1+< a_m^2 >)$ is approximately $< a_m^2 >$ for small perturbations.

If the b-function represents the signal component correctly to within an amplitude factor ε, b_j equals $\mu-\varepsilon s_j$, and

$$< F_m^*(f)\, F_m(\ln(f/b)) > = 4\mu < |F_m(\delta)|^2 > - 2\mu < F^*(\delta) > F_m \ln(\mu/b) \tag{19}$$

$$= \mu^{-1}[(1 - \varepsilon)\ |S_m|^2 + |N_m|^2]$$

Phase dependence is absent since the phase variations due to the signal are matched to those in the prior knowledge. If the amplitude is correctly estimated, ε equals unity and the cross-entropy derivatives on average depend on the noise power only, so minimization of cross-entropy entails a reduction of the noise component while maintaining the signal component. If the amplitude is incorrect, then some loss of signal power is expected, the amount depending on the degree of mismatch, so the minimization of cross-entropy is equivalent to a reduction of power representing the mismatch between the prior function and the data.

VII. ENTROPY DERIVATIVES FOR CORRELATED NOISE MODELS

In this section it is shown that for correlated noise, entropy has a phase dependency. Let U_j represent variables which are independent and uniformly distributed over (0,1). From these a family of noise models can be derived using the relationship

$$V_j = \alpha U_j + \beta U_{j-1}$$

where α is a parameter between 0 and $\frac{1}{2}$ and β represents $1-\alpha$.

The mean and variance of V_j are $\frac{1}{2}$ and $(\alpha^2 + \beta^2)/12$, respectively. The correlation at lag 1 is $\alpha\beta/(\alpha^2 + \beta^2)$ and zero for greater lags. Through straightforward integration, analytic expressions can be obtained for $< V_j \ln V_{j+k} >$, denoted by $q_k(\alpha)$, in order to calculate $< w_k(\alpha) >$. In particular,

$$q_0(\alpha) = -5/12 - (\beta^2\alpha^{-1}\ln\beta - \alpha^2\beta^{-1}\ln\alpha)/6 \;, \tag{20}$$

$$q_1(\alpha) = -3/4 + \alpha/\beta + \alpha\ln\alpha(6\alpha-3-\alpha^2)/12\beta^2 + \beta\ln\beta(\alpha+3)/12\alpha \;,$$

$$q_{N-1}(\alpha) + q_{-1}(\alpha) = -7/12 + (\beta+\alpha)/6\alpha + \alpha\ln\alpha(\alpha-4)/12\beta + \beta\ln\beta(2-4\alpha-\alpha^2)/12\alpha^2$$

and $q(\alpha) \equiv q_k(\alpha) = q_{N-k}(\alpha)$

$$= -3/4 - \alpha\ln\alpha/4\beta - \beta\ln\beta/4\alpha$$

for k not equal to 0 or 1.

Furthermore, for $m=1,2,\ldots,(M/2)-1$,

$$< F_m(w(\alpha)) > = MQ_0(\alpha) + (M-1)Q_1(\alpha)\cos 2\pi m/M + i(M-1)\,Q_2(\alpha)\sin 2\pi m/M \tag{21}$$

where $Q_0(\alpha) = q_1(\alpha) - q(\alpha)$,

$$Q_1(\alpha) = q_1(\alpha) + q_{N-1}(\alpha) - 2q(\alpha)$$

$$Q_2(\alpha) = q_{N-1}(\alpha) - q_1(\alpha) \;.$$

Q_0, Q_1 and Q_2 are positive for $0 < \alpha < \frac{1}{2}$, and Q_2 is small compared to the other two. Also $Q_2(\beta) = -Q_2(\alpha)$, whereas the other two functions are symmetric. The imaginary part depends upon $< x_j \ln x_{j+1} > - < x_{j+1} \ln x_j >$, which reflects asymmetry with regard to the spatial coordinate direction.

A related model for negatively correlated noise is given by

$$V_j = -\alpha U_j + \beta U_{j-1} + \alpha, \tag{22}$$

for which the correlation function of V_j at lag 1 is $-\alpha\beta/(\alpha^2 + \beta^2)$. It can be shown that

$$< F_m(w(-\alpha)) > = MQ_o(\alpha) - (M-1)Q_1(\alpha) \cos 2\pi m/M - i(M-1)Q_2 \sin 2\pi m/M \tag{23}$$

for $0 \leq \alpha \leq \frac{1}{2}$.

As a function of frequency the real part of $< F_m(w(\alpha)) >$ decreases with m whereas that of $< F_m(w(-\alpha)) >$ increases. From the relationship given in Eq. (12), it can be seen that to increase the entropy, larger decreases in amplitude are made at low frequencies than at high frequencies when the correlation is positive. The action is that of a high-pass filter. For negative correlation, the opposite applies, so the action is that of a low-pass filter. For uncorrelated noise, the derivatives are independent of frequency which indicates a uniform decrease in power is required over the flat spectrum.

In general the relationship between the auto-correlation function and the auto-cross-entropy function is reflected in a relationship between the power spectrum, and the entropy derivatives. If q_k equals $< f_j \ln f_{j+k} >$, then

$$< F_m(w) > = Mq_o + Mq_{M/2} \cos \pi m$$

$$+ \sum_{k-1}^{M/2-1} [(M-k)(q_k+q_k) + k(q_{M-k}+q_{k-M})] \cos 2\pi mk/M$$

$$+ i \sum_{k-1}^{M/2-1} [(M-k)(q_k-q_{-k} + k(q_{M-k}-q_{k-M})] \sin 2\pi mk/M. \tag{24}$$

If γ_k represents $< (f_j-\mu)(f_{j+k}-\mu) >$, where μ represents the mean of f, then the Fourier transform of $\sum_{k=0}^{M-1} < f_j-\mu)(f_{j+k}-\mu) >$ yields the power spectrum in the form

$$< |F_m(f)|^2 > = M\gamma_0 + M\gamma_{M/2}\cos\pi m + 2M \sum_{k=0}^{M-1} \gamma_k \cos 2\pi m/M \qquad (25)$$

This result may be obtained by replacing the q's in Eq. (24) by the corresponding γ's and reducing the result through application of the conditions of symmetry: $\gamma_k=\gamma_{N-k}=\gamma_{-k}$. A significant correlation at lag k, say, will be reflected in a relatively large value for the auto-cross-entropy, but the relationship is not simply stated because of the non-linearity introduced by the ln-function.

VIII. ILLUSTRATIVE EXAMPLES

In this section results are presented to illustrate the characteristics of signal reconstructions under various conditions of prior knowledge and constraints. The original data, the signal to be recovered, and various solutions are shown in Fig. 1. Because of its large steps, the function is not particularly well-suited to Fourier analysis, on the other hand, it is largely of low frequency content and is symmetric. The noise is additive Gaussian noise of zero mean and standard deviation equal to the size of the steps in the signal.

The signal is represented by one data set of 64 equispaced samples. Noise-only data were available for 56 sets of this size from which statistics of Fourier phase and magnitude were obtained for each frequency. This process was included since it simulated real situations in which noise samples are plentiful. It is assumed that the mean value of the signal is known.

For the maximum entropy solution, it is assumed that the signal is a constant, a condition of minimum prior knowledge, apart from the constraint that the signal be positive. For the minimum cross-entropy solutions, it is assumed that the signal is known exactly, so it is of the given

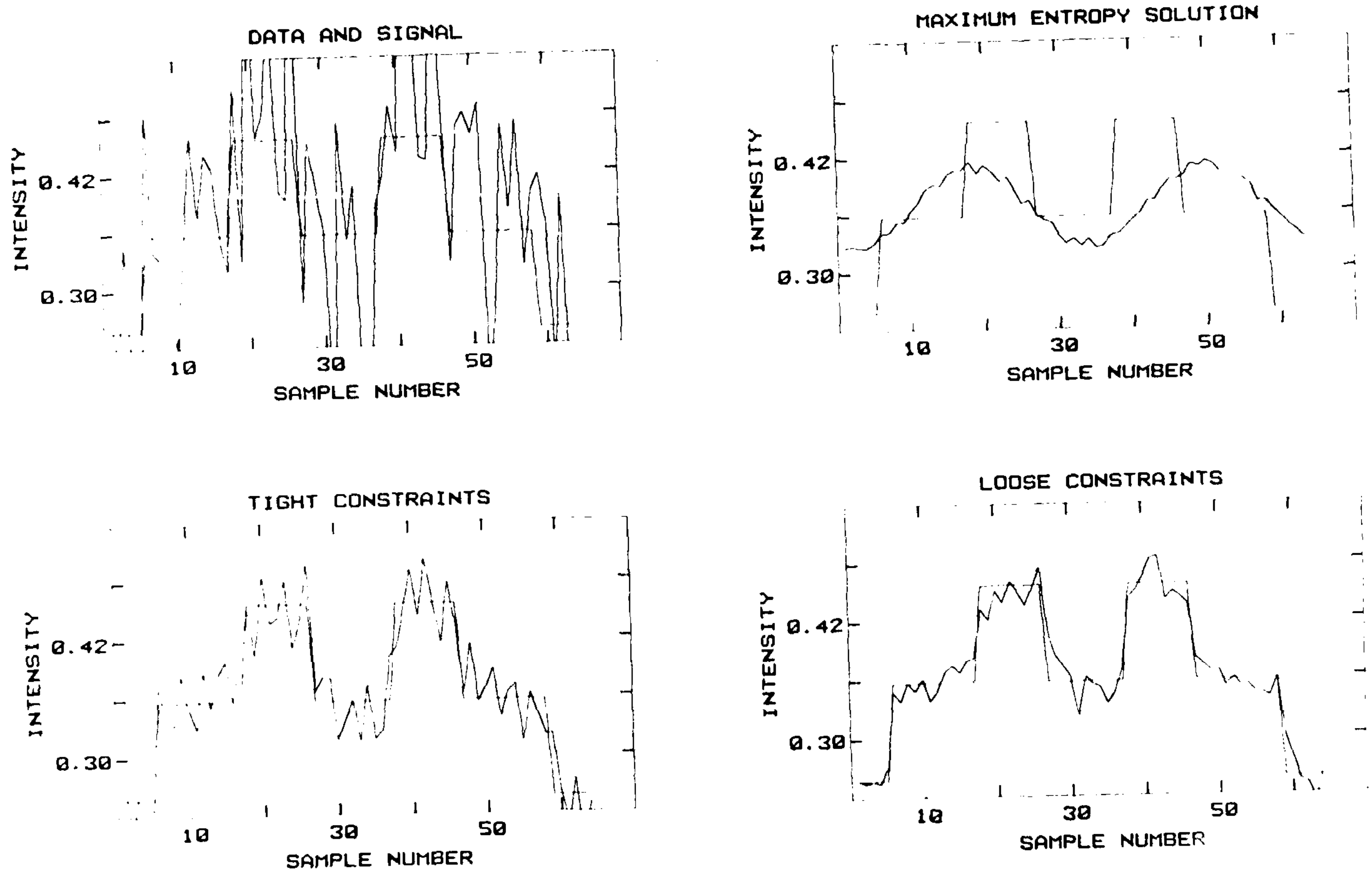

Figure 1. Plots of the original data and of signal representations obtained under various constraint and prior knowledge conditions. The true signal is shown in each plot.

form unless otherwise indicated by the data. The aim of the recovery process in these cases is not so much to obtain the assumed correct result as it is to evaluate differences between the expected signal and that produced from a particular data set.

The upper left quadrant of Fig. 1 shows the signal and the signal-plus-noise data which served as the signal representation. The upper right quadrant shows a maximum entropy solution with "loose" constraints, that is, constraints in the Fourier domain which were chosen to represent several standard deviations in the estimated noise power and phase. (A more exact definition is given below). The solution consists largely of a single low-frequency component. Sharp edges are not indicated in this "maximally smooth" solution.

The lower right quadrant of Fig. 1 shows the minimum cross-entropy solution for the same constraint bounds but with exact prior knowledge. The peaks and edges of the reconstruction match those of the signal. The standard deviation of the error is 0.017 and the maximum absolute error is 0.051 whereas for the original data these values were 0.114 and 0.319. The noise power has been reduced by a factor of 50.

The general appearance of this reconstruction is similar to that obtained from the Wiener filter (not shown) with exact knowledge of signal and noise spectra, except that the Wiener filter solution is slightly displaced to the right and lacks edge definition. For the Wiener filter solution the maximum absolute error is 0.058 and the standard deviation of the error is 0.031 which represents a reduction in noise power by a factor of 14. In this case the Wiener filter solution is a good initial estimate to the reconstruction for loose constraints. The main improvement accrues from phase shifts.

The lower left quadrant of Fig. 1 shows the effects of tightening the constraints in the Fourier domain so that more credence is afforded the data relative to the prior knowledge. In this case a high frequency component is evident. The standard deviation of the error is 0.059, which represents a noise power reduction factor of 3.7.

Generally, Fourier components are reduced by the minimization process to their lower bounds provided they are not otherwise constrained by the prior knowledge or restrictions in the spatial domain. For the "loose" constraints, the lower bound of the magnitude was zero unless the Fourier magnitude squared was 7 times the noise power standard deviation. Since the noise power at a given frequency is distributed as a chi-squared variable with mean equal to σ, the probability of a non-zero lower bound was roughly 1 out of 1000. For the "tight" constraints this probability was roughly 1 out of 50. The sample noise power for the high frequency which is evident in the solution for tight constraints was 6.8 times the noise mean. A peak of this magnitude is expected to occur roughly once in 1000 cases. At this frequency the tight constraints imposed a lower bound well above zero so that this component appears as a part of the solution. For loose constraints, the lower bound was zero and the contributions from this frequency were much reduced.

It may be thought that the data incorrectly indicate such a component as being part of the signal, but in practice a better view to take in this regard is that such a component is present at a certain level of confidence. If the analyst were to have prior knowledge with regard to the frequency composition of the signal, this knowledge should be incorporated into the constraints. If the signal was thought to have no isolated narrowband components, then the bounds on the Fourier magnitudes could incorporate several adjacent frequencies, perhaps through an averaging procedure, to reduce the standard deviation of the noise power estimates. On the other hand, if narrow-band components are not only possible but also of interest, then it would be incorrect to treat the components in groups in this manner. Flexibility in the imposition of constraints is a major benefit of treating data in the method proposed.

IX. CONCLUSIONS

It has been shown that a general purpose minimization method based on quasi-Newton search techniques can be applied to the problem of recovery of signals from noisy data when noise-only data are available for estimation of related statistics. A finite Fourier transformation is applied to the data and constraints applied to the Fourier coefficients consistent with the noise statistics. By requiring that the data be positive, the target function can be chosen to be the generalized entropy. The method then yields smooth solutions in the absense of prior knowledge (the maximum entropy solution) or solutions which tend to conform smoothly to the prior-knowledge solution (the minimum cross-entropy solution).

A major advantage of the method is the flexibility with which constraints may be imposed in either the frequency domain or the temporal (or spatial) domain. This allows for direct application of prior information to the process of solution. A second major advantage is the adaptability of the method with regard to the degree of credibility assigned to the signal data. Tight bounds on the constraints yield solutions which conform closely to the data. Loose bounds yields solutions which resemble closely the expected results. Thus the analyst has a full range of solutions from which to choose that which is most appropriate for the particular problem.

The selection of the appropriate constraint levels may be considered as a problem in statistical analysis and levels of confidence may be applied to solutions. This was demonstrated by an illustrative example.

The minimization process requires the derivatives of the generalized entropy with respect to the Fourier coefficients. These were obtained analytically and results were studied for two cases of interest. It was shown that phase is important not only for cross-entropy solutions but also for maximum entropy solutions when the noise is correlated. The analytical results indicate that the minimization process tends to produce solutions satisfying lower bounds for the power spectral coefficients. Thus it is expected that solutions will resemble those obtained by the simple noise subtraction method given by Eq. (8) unless otherwise constrained.

The examples given pertained to one-dimensional variables but the results can be extended to images [1]. A major disadvantage of the method is the computational time required. For this reason studies are proposed on special purpose minimization processes applicable to a more restrictive class of signal, for example, severely band-limited signals.

REFERENCES

[1] Barrodale I., C.A. Zala and R.F. MacKinnon, "Image processing by a maximum entropy procedure incorporating frequency domain bounds and prior knowledge" in Conference Proceedings of IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, June 4-4, 1987, Victoria, B.C. (IEEE Publishing Services, New York, 1987).

[2] Barakat Richard, and Garry Newsam, :Algorithms for reconstruction of partially known, band-limited Fourier-transform pairs from noisy data", J. Opt. Soc. Amer., A, vol. 2, pp. 2027-2039, 1985.

[3] Jaynes Edwin T., "On the rationale of maximum entropy methods", Proc. IEEE, vol. 70, pp. 939-952, 1982.

[4] Frieden B.R., "Image enhancement and restoration" from Picture Processing and Digital Filtering, T.S. Huang (editor), Topics in Applied Physics, vol. 6. (Springer-Verlag, New York, 1979).

[5] Skilling John, "The maximum entropy method", Nature, vol. 309, pp. 748-749, 1984.

[6] Shore John E. and Rodney W. Johnson, "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy", IEEE Trans. Inform. Theory, vol. IT-26, pp. 26-37, 1980.

[7] Frieden B. Roy, Probability, Statistical Optics and Data Testing, (Springer-Verlag, New York, 1983).

[8] Jenkins Gwilym M. and Donald G. Watts, Spectral Analysis, (Holden-Day, San Francisco, 1968).

[9] Gonzalez Rafael C. and Paul Wintz, Digital Image Processing, (Addison-Wesley, Reading, Mass., 1977).

[10] Hunt B.R., "The application of constrained least squares estimation to image restoration by digital computer", IEEE Trans. Computers, vol. C-22, pp. 805-812, 1973.

[11] Gull S.F. and G.J. Daniell, "Image reconstruction from incomplete and noisy data", Nature, vol. 272, pp. 686-690, 1978.

[12] Lim Jae S., "Image restoration by short space spectral subtraction", IEEE Trans. Acoust., Speech, Signal Procesing, vol. ASSP-28, pp.191-197, 1980.

[13] Whalen Anthony D., Detection of Signals in Noise, pp. 113-116, (Academic Press, New York, 1971).

[14] Oppenheim Alan V. and Jae S. Lim, "The importance of phase in signals", Proc. IEEE, vol. 69, pp. 529-541, 1981.

[15] Fienup J.R., "Comments on 'The reconstruction of a multidimensional sequence from the phase or magnitude of its Fourier transform'", IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-31, 738-739, 1983.

[16] Curtis Susan R., Alan V. Oppenheim and Jae S. Lim, "Signal reconstruction from Fourier transform sign information", IEEE Trans. Acoust. Speech, Signal Processing, vol. ASSP-33, pp. 643-657, 1985.

[17] NAG Fortran Library, algorithm E04VAF based on methods described by W. Murray, "Methods for constrained optimization", Ch. 12 of L.C.W. Dixon (editor), Optimization in Action, (Academic Press, 1967], National Algorithms group, 1101 31st Street, Suite 100, Downers Grove, Illinois 60515, USA.

REPORT NO.: DREP Technical Memorandum 88-9

TITLE: MAXIMUM ENTROPY CRITERIA APPLIED TO SIGNAL RECOVERY

AUTHORS: Robert MacKinnon and Michael J. Wilmut

DATED:

SECURITY GRADING: UNCLASSIFIED

3 - DSIS
Circulate to:
DRDM
DSP

1 - DREV
1 - DREO
1 - DREA

1 - DGAEM
1 - DAR
1 - DGMEM
1 - DGIS/DSTI
2 - ORAE Library
1 - D Met Oc
1 - Maritime Tech Library
1 - Bedford Inst of Oceanography
Library
1 - NRC/CISTI (Microfiche only)
1 - RRMC Dept of Oceanography
1 - RMC
1 - CFMWS
1 - CRC Library

1 - CDLS(L) CDR
1 - CDLS(W) CDR
1 - DRA Paris

1 - Commander, Maritime Command
Attn: MC/ORD
1 - SSO OpRsch/MARPAC
1 - Pacific Environment Institute
1 - Great Lakes Inst,
University of Toronto

1 - RADARSAT Office
110 O'Connor St., Suite 200
Ottawa, Ontario K1P 5M9
Attn: Mr. Ed Langham

1 - RESORS
Canada Centre Remote Sensing
240 Bank Street, 5th Floor
Ottawa, Ontario K1A 0Y7
Attn: Mr. G. Thirlwall

BRITAIN

Ministry of Defence
1 - DRIC

UNITED STATES

3 - DTIC
1 - Naval Research Laboratory
Washington, D.C. 20375
1 - Environmental Research Institute
of Michigan
P.O. Box 8618
Ann Arbor, Michigan 48107
Attn: Dr. R. Shuchman

UNCLASSIFIED

SECURITY CLASSIFICATION OF FORM
(highest classification of Title, Abstract, Keywords)

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Establishment sponsoring a contractor's report, or tasking agency, are entered in section 8.)

DREP

2. SECURITY CLASSIFICATION (overall security classification of the document, including special warning terms if applicable)

UNCLASSIFIED

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S,C,R or U) in parentheses after the title.)

MAXIMUM ENTROPY CRITERIA APPLIED TO SIGNAL RECOVERY (U).

4. AUTHORS (Last name, first name, middle initial)

MACKINNON, ROBERT F. AND WILMUT, MICHAEL J.

5. DATE OF PUBLICATION (month and year of publication of document)

6a. NO. OF PAGES (total containing information. Include Annexes, Appendices, etc.)

23

6b. NO. OF REFS (total cited in document)

17

7. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

TECHNICAL MEMORANDUM

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address.)

DREP

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant)

DRDM-04

9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written)

N/A

10a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.)

10b. OTHER DOCUMENT NOS. (Any other numbers which may be assigned this document either by the originator or by the sponsor)

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

(X) Unlimited distribution
( ) Distribution limited to defence departments and defence contractors; further distribution only as approved
( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
( ) Distribution limited to government departments and agencies; further distribution only as approved
( ) Distribution limited to defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected.)

UNCLASSIFIED

SECURITY CLASSIFICATION OF FORM

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both offical languages unless the text is bilingual).

A method based on the minimization of cross-entropy is presented for the recovery of signals from noisy data either in the form of time series or images. Finite Fourier transforms are applied to the data and constraints are placed on the magnitude and phase of the Fourier coefficients based on their statistics for noise-only data. The minimization of cross-entropy is achieved through application of well-established functional minimization techniques which allow for further constraints in the spatial, temporal or frequency domain. Derivatives of the entropy function are obtained analytically and the results applied to the cases of correlated noise and of signal perturbations about a mean. Demonstrations of applications to one-dimensional data are presented.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

image processing
maximum entropy methods
signal recovery
spatial analysis
signal processing