

Parameter Estimation and Extraction of Helicopter Signals Observed with a Wide-Band Interference

Meir Feder, *Member, IEEE*

Abstract—The acoustical signal generated by a helicopter, as well as many other signals, e.g., a voiced speech signal, can be described as periodic or almost periodic. When these signals are observed in the presence of other additive wide-band signals, it becomes interesting to separate these two kinds of signals. In this paper we specifically present an iterative method for separating helicopter signals from wide-band interference modeled as an autoregressive (AR) process. The suggested algorithms can also be used for other harmonic signals. Experimental results of our algorithm using real helicopter signals are presented to demonstrate the practical applicability of the method.

I. INTRODUCTION

MANY signals observed in a variety of applications can be characterized as periodic or almost periodic. A voiced speech signal and the acoustic signal generated by a helicopter are examples for such signals. These signals are usually characterized in the frequency domain by their harmonic spectrum whose fundamental frequency is the inverse of the period. In the time domain, the peaks of the autocorrelation function of such signal indicate as well the period of these signals.

A basic problem concerning these signals is the estimation of the period or the fundamental frequency. Many methods have been suggested for solving this problem. In the speech context these methods are referred to as pitch detection, e.g., [1]–[4]. Other methods have been suggested in the helicopter context, e.g., [5], [6]. In a general signal processing context, methods were suggested in [7] for the parameter estimation problem and in [8] for the more general estimation and signal enhancement problem. In this paper we consider the problem of estimating the periodic signal parameters in the presence of an interference. This is an interesting problem since most of the previously suggested methods often show unacceptable degradation in performance in the presence of an interference characterized as colored noise. A related problem to be considered in this noisy case is the estimation of the

Manuscript received August 23, 1988; revised September 5, 1991. This work was supported in part by the Machine Intelligence Technology Group at M.I.T. Lincoln Laboratory, which is sponsored by the U.S. Air Force; in part by the Advanced Research Projects Agency monitored by ONR under Contract N00014-81-K-0742; and in part by the National Science Foundation under Grant ECS-8407285.

The author is with the Department of Electrical Engineering-Systems, Faculty of Engineering, Tel-Aviv University, Tel-Aviv, 69978, Israel.

IEEE Log Number 9203340.

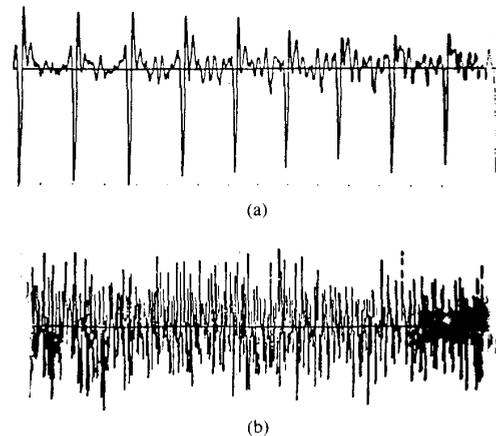


Fig. 1. (a) Typical helicopter signal. (b) Typical jet plane signal.

periodic signal itself—either for enhancing it or for canceling it out.

The motivation for this problem comes from a real problem where a helicopter acoustic signal is observed with an acoustic jet plane signal. Both signals may be of interest and it is desired to separate them for further processing. Typical examples of the helicopter signal waveform and the jet plane waveform are given in Fig. 1.

A model-based approach is used in this paper. The periodic (or close to periodic) signal is defined up to some parameters, one of them is T , the period; and the others describe the deviation from strict periodicity. The wide-band signal is modeled as a Gaussian random process, with unknown spectrum. To simplify the analysis, we assume an AR model whose parameters determine the bandwidth and the center frequency of this signal.

The likelihood of the observed signal given all the model parameters can be maximized to produce an estimate of these parameters. However, this maximization is complicated. We suggest an iterative algorithm for maximizing the resulting likelihood function. This algorithm is closely related to the algorithms described in [9], and can be described as an implementation of the general expectation-maximization (EM) algorithm [10].

The paper is organized as follows. Section II will be devoted to the detailed analysis of the periodic or the almost periodic signal. In Section III we present the problem of separating almost periodic signals from wide-band

signals, and derive an iterative algorithm that accomplishes this separation and estimates the various model parameters. We also discuss a second version of the algorithm to the case in which the almost periodic signal has two fundamental frequencies. In Section IV we present experimental results with the suggested algorithms. The necessary background, i.e., a short description of the EM algorithms is provided in the Appendix.

II. WAVEFORM AND PARAMETER ESTIMATION OF ALMOST PERIODIC SIGNALS

In this section we investigate methods for waveform and parameter estimation of almost periodic signals. At this point we do not define exactly what we mean by "almost" periodic; it will be clear in the various cases later in the paper. However, we can say in general that an almost periodic signal $x(t)$ satisfies

$$x(t) \approx x(t + T) \quad (1)$$

for some period T , at least along some time window.

Throughout this section we assume that we observe the almost periodic signal $x(t)$ with an additive error signal $e(t)$, i.e., we observe $y(t) = x(t) + e(t)$. This error signal accommodates for measurements errors, deviations from the model, etc. We start by examining the case where $x(t)$ is strictly periodic, and then investigate possible deviations of $x(t)$ from strict periodicity. The strictly periodic case is the simplest, but the algorithms derived for it will be the basis for the more general cases in which deviations for strict periodicity will be allowed.

A. Strictly Periodic Signal

A strictly periodic signal, $x(t)$, whose period is T , is a signal that satisfies the equality

$$x(t) = x(t + T) \quad \forall T \quad (2)$$

which implies that

$$x(t) = x(t + kT) \quad k = 0, \pm 1, \pm 2, \dots \quad (3)$$

We define the period waveform $a(t)$ as

$$a(t) = \begin{cases} x(t) & 0 \leq t < T \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

A different way to describe this signal is as a convolution, i.e.,

$$x(t) = a(t) * p(t) \quad (5)$$

where $p(t)$ is an impulse train

$$p(t) = \sum_k \delta(t - kT) \quad (6)$$

which leads to a modulated impulse train behavior in the frequency domain.

Let us assume, first, that the period T is given. Consider the samples of the observed signal at T time units

apart, i.e., $\{y(t + kT)\}$. In order to estimate a sample of the period waveform $a(t)$, define the vector $\mathbf{y}(t)$ as

$$\mathbf{y}(t) = [y(t + K_1 T), y(t + (K_1 + 1)T), \dots, y(t + (K_2 - 1)T), y(t + K_2 T)]^T \quad 0 \leq t < T \quad (7)$$

where $K_1(K_2)$ is the smallest (largest) k such that $t + kT$ is in the observation window. Usually, for symmetry, $K_1 = -K_2$. From (3), (4), and (7),

$$\mathbf{y}(t) = \mathbf{1}a(t) + \mathbf{e}(t) \quad (8)$$

where $\mathbf{1} = [1, \dots, 1]^T$ and $\mathbf{e}(t)$ is defined as

$$\mathbf{e}(t) = [e(t + K_1 T), e(t + (K_1 + 1)T), \dots, e(t + (K_2 - 1)T), e(t + K_2 T)]^T \quad 0 \leq t < T. \quad (9)$$

If we further assume that the signal $e(t)$ is white, the vectors $\mathbf{y}(t)$ are independent at different time points. Thus, we may process each time sample independently. If $e(t)$ is also assumed to be Gaussian, the vector $\mathbf{e}(t)$ is zero-mean Gaussian with covariance matrix $\sigma^2 \mathbf{I}$ and so the ML estimator of $a(t)$ is the solution to the least squares problem $\min \|\mathbf{y}(t) - \mathbf{1}a(t)\|^2$, which is a simple average

$$\hat{a}(t) = \frac{1}{K} \mathbf{1}^T \mathbf{e}(t) = \frac{1}{K} \sum_{k=K_1}^{K_2} y(t + kT) \quad 0 \leq t < T \quad (10)$$

where $K = K_2 - K_1 + 1$.

When the error samples are correlated, with a correlation matrix $R_e(t) = E\{\mathbf{e}(t)\mathbf{e}(t)^T\}$, the ML criterion is equivalent to a weighted least squares criterion, i.e., minimizing $[\mathbf{y}(t) - \mathbf{1}a(t)]^T R_e^{-1} [\mathbf{y}(t) - \mathbf{1}a(t)]$, and we get

$$\begin{aligned} \hat{a}(t) &= [(\mathbf{1}^T R_e^{-1}(t) \mathbf{1})]^{-1} \mathbf{1}^T R_e^{-1}(t) \mathbf{y}(t) \\ &= \frac{\sum_{k=K_1}^{K_2} \alpha_k(t) y(t + kT)}{\sum_{k=K_1}^{K_2} \alpha_k(t)} \quad 0 \leq t < T \end{aligned} \quad (11)$$

where $\alpha_k(t)$ is the k th coefficient of the vector $\mathbf{1}^T R_e^{-1}(t)$. The estimate in (11) is achieved by processing each time sample separately; thus, it is optimal only if the error signal samples at different time points are uncorrelated, i.e., $E\{e(t)e(\sigma)\} = 0$, $t \neq \sigma$, $0 \leq t, \sigma < T$. We note that while the estimate of $a(t)$ in (10) is independent of the exact knowledge of the statistical properties of $e(t)$, i.e., of σ^2 , we need to know R_e , i.e., the correlation (or the spectrum) of $e(t)$, in order to obtain the estimate of (11).

The estimate of $a(t)$ in (11) can be regarded as a weighted average over the samples of the observed signal in the various periods. It is interesting to note that we get a similar, although slightly different weighted average solution for the problem of estimating $a(t)$ using a model

that deviates from the periodic model. Suppose that $x(t)$ satisfies

$$x(t) = w_k(t) \cdot x(t + kT). \quad (12)$$

The signal $x(t)$ will be considered almost periodic if $w_k(t) \approx 1$, and it is a smooth function. One immediately observes that when the coefficients of $e(t)$ are uncorrelated, the estimate of the periodic waveform $a(t)$ is given by

$$\hat{a}(t) = \frac{\sum_{k=K_1}^{K_2} w_k(t) y(t + kT)}{\sum_{k=K_1}^{K_2} w_k^2(t)}. \quad (13)$$

Again, in order to use this estimator one has to know the "modulating" waveform $w_k(t)$. The simultaneous estimation of $w_k(t)$ (with some assumptions) and the periodic waveform $a(t)$, is given in Appendix A.

So far no *a priori* knowledge and no assumptions about the periodic waveform $a(t)$ have been used. The solutions of (10), (11), and (13) were obtained independently at each time point as the unconstrained solutions of the appropriate optimization problems. We may have some deterministic knowledge about $a(t)$, e.g., that it is smooth and its derivative cannot exceed a certain value. However, usually it is hard to incorporate these constraints. We will thus try to express our knowledge by assuming *a priori* statistical properties for $a(t)$.

Since the signal is periodic with period T , we may consider its Fourier series coefficients, which are given by

$$A(\omega_l) = \frac{1}{\sqrt{T}} \int_0^T a(t) e^{-j\omega_l t} dt \quad \omega_l = l \frac{2\pi}{T}. \quad (14)$$

We assume that, *a priori*, these Fourier coefficients are zero-mean Gaussian complex random variables, with variance,

$$E\{A(\omega_l)A^*(\omega_l)\} = S_a(\omega_l). \quad (15)$$

We also assume that the Fourier series coefficient, at different frequencies, are uncorrelated. The periodic waveform is related to the Fourier coefficients by the inverse formula

$$a(t) = \frac{1}{\sqrt{T}} \sum_l A(\omega_l) e^{j\omega_l t}. \quad (16)$$

We note that the presentation of (15), together with the assumptions on $A(\omega_k)$, is Rice's presentation of periodic Gaussian signals, see [11] and [12].

For estimating the Fourier coefficients, consider the Fourier series of the successive blocks of size T of the observed signal, i.e.,

$$Y_k(\omega_l) = \frac{1}{\sqrt{T}} \int_0^T y(t + kT) e^{-j\omega_l t} dt = A(\omega_l) + E_k(\omega_l) \quad (17)$$

where $E_k(\omega_l)$ is the Fourier series coefficient of the k th block of the error signal. Using a vector notation, define

$$\mathbf{Y}(\omega_l) = [Y_{K_1}(\omega_l), \dots, Y_{K_2}(\omega_l)]$$

and define similarly $\mathbf{E}(\omega_l)$, so

$$\mathbf{Y}(\omega_l) = \mathbf{1}^T \mathbf{A}(\omega_l) + \mathbf{E}(\omega_l). \quad (18)$$

For simplicity, assume again that the signal $e(t)$ is white, i.e., the components of $\mathbf{E}(\omega_l)$ are uncorrelated with covariance $\sigma_e^2 \mathbf{I}$, as are the various vectors $\mathbf{E}(\cdot)$ at different frequencies. In this case the signal's Fourier series coefficients, $A(\omega_l)$, can be estimated separately at each frequency. The solution to the MAP problem in the Gaussian case which is the optimal linear estimation is given by

$$\hat{A}(\omega_l) = S_a(\omega_l) \mathbf{1}^T [\mathbf{1} S_a(\omega_l) \mathbf{1}^T + \sigma_e^2 \mathbf{I}]^{-1} \cdot \mathbf{Y}(\omega_l). \quad (19)$$

Note that the solution in (19) depends on the knowledge of the spectrum $S_a(\omega)$.

Estimating the Period: For each value of the period T we can find an estimate of the period waveform $\hat{a}(t; T)$ depending on our *a priori* knowledge and assumptions about $e(t)$, using one of the methods described above. We can then reconstruct an estimate of the periodic signal $x(t)$ and the error signal $\hat{e}(t; T) = y(t) - \hat{x}(t; T)$ based on the estimate of the period waveform $\hat{a}(t; T)$. The period can then be estimated by minimizing the resulting goal function with respect to T .

Following the discussion above, when it is assumed that $e(t)$ is white and Gaussian the period will be estimated by minimizing

$$\begin{aligned} L(T) &= \frac{1}{KT} \int_{\mathfrak{J}} (y(t) - \hat{x}(t; T))^2 dt \\ &= \frac{1}{KT} \sum_{k=K_1}^{K_2} \int_0^T (y(t + kT) - \hat{a}(t; T))^2 dt \end{aligned} \quad (20)$$

where \mathfrak{J} is the observation window, and where $\hat{a}(t; T)$ is given by (10). Recall that (10) is an average of the observed signal over samples which are T time apart. Substituting this estimate in (20) we can immediately observe the relation between minimizing (20) and searching for the peaks of the autocorrelation function of $y(t)$.

If $e(t)$ is not assumed to be white, but we know its correlation function $R_e(t, s)$ the criterion for estimating T will be

$$\begin{aligned} L(T) &= \frac{1}{KT} \int_{\mathfrak{J}} \int_{\mathfrak{J}} (y(t) - \hat{x}(t; T)) \\ &\quad \cdot R_e(t, s) (y(s) - \hat{x}(s; T)) dt ds \end{aligned} \quad (21)$$

where now we have to use the estimate of $\hat{a}(t; T)$ given by (11).

If we assume that each period of $x(t)$ is modulated by $w_k(t)$ as in (12), we have to include these weights when we generate $\hat{x}(t; T)$ from the period waveform $\hat{a}(t; T)$ estimated by (13).

For the case where we have statistical *a priori* knowledge on $a(t)$ in the form presented by (14) and (15), a term that represents the *a priori* knowledge, i.e.,

$$\sum_{\omega_k} \log \det S_a(\omega_k) + A(\omega_k; T)^* S_a(\omega_k) A(\omega_k; T) \quad (22)$$

where $*$ denotes the complex conjugate, must be added to (20). T is estimated by minimizing the sum of both terms.

The approach presented so far can be regarded as a generalization of the ML pitch estimation method suggested in [2]. More importantly, it sets the framework for the results presented below which consider deviations from strict periodicity.

B. Deviations from the Strictly Periodic Model

In reality, signals which are strictly periodic in the entire observation window are, unfortunately, rare. For example, the fundamental frequency of a voiced speech signal (the pitch) can change in time. Similarly, the acoustic helicopter signal is not strictly periodic due to Doppler shift effects. Thus our model and processing procedures must accommodate some deviations from strict periodicity. Note that above, and in Appendix A, we have already discussed one deviation from a strict periodicity by allowing a different gain w_k for each period. However, the more interesting deviations from the periodic model are those where the "period" T is not constant.

In this section we consider two such models. In the first we assume that the period or the fundamental frequency is changing slowly, so that it may be considered constant in a short enough interval. If the changes in the fundamental frequency are the result of a Doppler shift this assumption corresponds to a constant radial velocity in that interval. In the second model we assume that the period time T can change according to some parameter; thus, if the changes in the fundamental frequency are the result of a Doppler shift, this corresponds to assuming that the radial velocity is varying in time according to some parameters.

1) *Slowly Varying Period*: We assume that along a window of several time periods assumption is approximately valid, i.e., the signal part $x(t)$ satisfies

$$x(t) \approx x(t + kT) \quad k = \pm 1, \dots, \pm K \quad (23)$$

where K is a small enough number.

Given the period T , in a small enough window around time t , the estimate of the periodic waveform $a(t)$, say, according to the least squares goal function, is similar to (10) and is given by

$$\hat{a}(t; T) = \frac{1}{2K + 1} \sum_{-K}^K y(t + kT). \quad (24)$$

Thus, the "local" period can be estimated by minimizing

$$L(T) = \frac{1}{2\Delta} \int_{t-\Delta}^{t+\Delta} (y(t) - \hat{a}(t; T))^2 dt. \quad (25)$$

We note that if the period is constant in time, using as large as possible Δ will improve the performance. If the

integral in (23) is performed over the entire observation window we get the goal function of (20). However, since the period is slowly changing in time, a smaller Δ , usually on the order of T , is recommended.

Estimating a new period T , at each time point t , may be too expensive and unnecessary. We can use the same period estimate T along some window, and estimate the periodic waveform along that window according to (24). The period estimate will be updated by searching (25) at the rate we expect it to change.

The estimate of the periodic signal by (24) can be motivated by observing that (24) represents a convolution between the observed $y(t)$ signal and

$$p(t) = \sum_{-K}^K \frac{1}{2K + 1} \delta(t - kT). \quad (26)$$

This is a filter whose frequency response has peaks at frequencies $1/T$ apart and each peak has a bandwidth of $1/(2K + 1)T$. This filter is known as "comb filter." The shape of this filter can be modified slightly (e.g., can have a better sidelobe behavior) by using nonequal weights α_k in (26). The frequency response of such a filter with arbitrary α_k 's is given in Fig. 2. It is clear that this filter is aimed to pass only the periodic components, of period T , of the observed signal.

The goal of the suggested algorithm is to tune the comb filter to the observed signal. Adaptive comb filters have been suggested, e.g., [13], [4]. We have suggested here an explicit filter where the period estimate comes from the ML procedure derived above (which is similar to [2]). We also note that a different approach, based on the polynomial structure of the periodic signal with a finite number of harmonics has been suggested in [8]. Our approach is more direct, and as will be seen below, it can be used in the more complicated problem of separating the almost periodic signal and the wide-band interference signal.

2) *Parameterically Varying Period*: The direct approach for adapting the comb filter can be extended to the case where the period variation is controlled by some parameter. Recall that the waveform estimate of the periodic signal was possible since we could associate the signal $x(t)$ at each time point to a set of signal values at other time points $x(t + kT)$ which, due to periodicity, have equal values. This situation can be generalized as follows. Suppose that the desired signal satisfies

$$x(t_0(t)) = x(t_k(t)) \quad k = \pm 1, \pm 2, \dots \quad (27)$$

When the signal is strictly periodic we have, of course, $t_k(t) = t + kT$. However, the more general setting of (27) enables us to consider nonperiodic signals.

Suppose the time delay functions $t_k(t)$ depend on some parameters. One parameter will be T , the length or the period of the basic waveform $a(t)$ to be estimated. Other parameters define the deviation from the periodic assumptions. For example, suppose that due to Doppler effects, the original time axis of the periodic signal is compressed or stretched linearly in time at a rate α . This situation

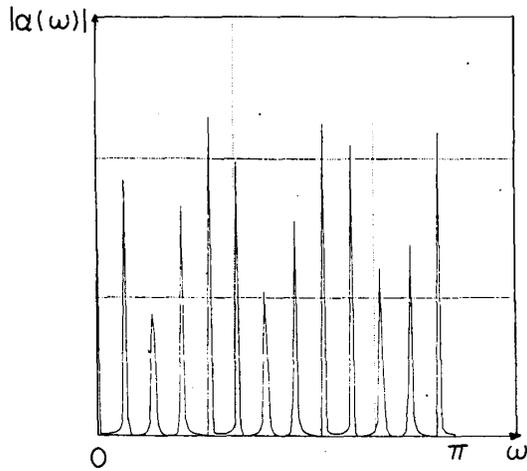


Fig. 2. Frequency response of a comb filter.

occurs when there is a radial acceleration and thus the radial velocity changes as αt . In this case the original periodic signal $x(t)$ is modified and becomes

$$x^{(\text{dop})}(t) = x([1 + \alpha t]t). \quad (28)$$

We notice that in this case $t_k(t) = [1 + \alpha(t + kT)](t + kT)$, i.e., determined by two parameters α and T .

For each value of the parameters we can find the best estimate of the basic waveform $a(t)$ by using one of the goal functions and the appropriate processing procedures as presented in the previous section. We have, of course, to use the appropriate time index. Thus, using the least squares goal function, the estimate of $a(t)$ which is analogous to (10), is given by

$$\hat{a}(t_0(t)) = \frac{1}{K} \sum_{k=K_1}^{K_2} y(t_k(t)) \quad 0 \leq t \leq T \quad (29)$$

where $t_k(t) = [1 + \alpha(t + kT)](t + kT)$. As before, we fix $a(t_0(t))$ to be zero for $t < 0$ or $t > T$.

The unknown parameters that define the period and its variations can be estimated, as before, by substituting the waveform estimate (which is a function of these parameters) into the desired goal function for these parameters, and searching for the optimal parameter values. Thus, if we denote the solution to (29) by $\hat{a}(\cdot; T, \alpha)$, and we use again the least squares criterion, the suggested procedure for estimating α and T is given by

$$\hat{\alpha}, \hat{T} = \arg \min_{\alpha, T} \int \left(y([1 + \alpha t]t) - \sum_k \hat{a}(t_k(t); T, \alpha) \right)^2 dt. \quad (30)$$

The bounds of the integral in (30) as well as the choice of K_1 and K_2 in (29) (i.e., the number of "periods" we average on), can be determined by our *a priori* assumption about the size of the window for which we can assume that α and T are constants.

Substituting (29) into (30) leads to an expression containing quadratic terms of the observed signal $y(t)$. Thus, as expected, the goal function of (30) is strongly related to the autocorrelation of $y(t)$.

III. PARAMETER ESTIMATION AND SEPARATION OF ALMOST PERIODIC SIGNAL AND AR PROCESS

The main problem considered in this paper is the separation of a wide-band signal and a periodic (or almost periodic) signal. As noted above, the practical problem that motivated this research comes from a situation where we observe a helicopter acoustic signal together with a jet plane signal. In the sequel, we assume that the observed signal $y(t)$ can be written as

$$y(t) = h(t) + j(t) \quad (31)$$

where the signal $h(t)$, the helicopter signal, is composed of a signal $x(t)$ and a signal $e(t)$, and $x(t)$ is either strictly periodic or deviate from periodicity according to the models discussed above. To simplify the exposition, throughout this section it is further assumed that $e(t)$ is Gaussian and white. The signal $j(t)$, which represents the wide-band interference, is assumed to be a zero-mean stationary random process, independent of $h(t)$, whose spectrum depends on some unknown parameters. A possible model for $j(t)$ is a Gaussian AR process. In this case the power spectrum of $j(t)$ is given by

$$S_j(\omega) = \frac{G}{\left| 1 - \sum_{i=1}^p a_i e^{-j\omega i} \right|^2} \quad (32)$$

where a_i are the AR parameters and p is the model order. A low model order is usually sufficient to capture the spectral shape of $j(t)$. The spectral parameters of $j(t)$ will be denoted ϕ .

The problem considered here is basically different, and may be substantially more complicated, than the problem of enhancement and parameter estimation of harmonic signal in a white noise. One reason for that lies in the fact that the AR process may, itself, show a narrow-band behavior when the poles are closed to the unit circle. Indeed, the previously proposed algorithms which work well in the white noise case will not perform well in our case. Nevertheless, the problem can still be solved and it is not ill-posed, since the noise waveform, even when its band becomes narrow, does not have the deterministic harmonic structure assumed for the (almost) periodic signal, and so these two signals can be separated. The algorithm derived in this section provides such a solution.

The likelihood of the observed data with respect to all the unknown parameters is derived in Appendix B. Maximizing this likelihood directly is complicated. Now, we know that if we observe only the signal $h(t)$, which contains the almost periodic signal and the white error signal, we can find the ML estimate of the parameters based on the techniques of Section II. Also, if we observe only the signal $j(t)$, the AR process, we can estimate its parame-

ters easily by solving the appropriate normal equations. We are now going to use these facts to derive an iterative algorithm based on the *EM* method, for estimating all parameters jointly.

We recall that for deriving algorithms based on the *EM* method we consider the observations as "incomplete" with respect to a more convenient set of data, called the "complete data." Each iteration has two steps, the *E* and the *M* steps. Denoting the complete data z , the observed data y , and the current value of the parameters $\theta^{(n)}$, the *E* step calculates

$$Q(\theta, \theta^{(n)}) = E \{ \log p(z) | y, \theta^{(n)} \} \quad (33)$$

and the *M* step calculates

$$\theta^{(n+1)} = \arg \max_{\theta} Q(\theta, \theta^{(n)}) \quad (34)$$

For exponential families of distributions, the *E* step estimates the sufficient statistics of the complete data by a conditional expectation, given the observations and the previous value of the parameters. The *M* step maximizes the likelihood of the complete data using the estimated statistics of the complete data instead of the true one which is unobserved. In Appendix C we discuss the *EM* method further. The iterations (33) and (34) are the generic form of the method. An explicit algorithm for our problem is derived below.

A. Derivation of the Iterative Procedure

Following the discussion above, there is natural choice of complete data—the signals $h(t)$ and $j(t)$ separated. Indeed, if $h(t)$ and $j(t)$ were observed separately, estimating the unknown parameters would have been simple. Being more specific, suppose the complete data is given. Since $h(t)$ and $j(t)$ are statistically independent, the log-likelihood of the complete data is the sum of the log-likelihood of $h(t)$ and $j(t)$, i.e.,

$$\log p(h(t), j(t)) = \underbrace{\log p(h(t))}_I + \underbrace{\log p(j(t))}_{II} \quad (35)$$

Now the first term depends only on the parameters of the (almost) periodic signal, say T and $a(t)$. The second term depends only on the parameters ϕ of $j(t)$. Thus each term can be maximized separately with respect to the appropriate variables.

The maximization of term *I* above was discussed in the previous section. For example, when the signal is strictly periodic and $e(t)$ is Gaussian and white we minimize

$$\hat{T} = \arg \min_T L(T) = \arg \min_T \sum_t (h(t) - \hat{a}(t; T))^2 \quad (36)$$

where $\hat{a}(t; T)$ is a periodic signal for which each period waveform is given by

$$\hat{a}(t; T) = \frac{1}{K} \sum_{k=k_1}^{K_2} h(t + kT) \quad 0 \leq t < T. \quad (37)$$

Note that the observed signal is assumed to be discrete time; however, the period T may not be an integer and we

may have to use interpolation when $t + kT$ is not an integer. When the signal deviates from strict periodicity, we use instead formulas analogous to (30) and (29), where $h(t)$ replaces $y(t)$, and summation over t replaces the integration. In any case, the statistics of $h(t)$ needed for estimating its parameters contain linear terms of $h(t)$ and quadratic terms of the form $h(t)h(\sigma)$.

The term *II* of (35), which depends only on the spectral parameters of $j(t)$, can be easily maximized for the case when $j(t)$ is an AR process. It is well known that in this case maximizing the likelihood is (approximately) equivalent to minimizing the prediction error of the AR process i.e., minimizing

$$\epsilon^2 = \sum_t \left(j(t) - \sum_{k=1}^p a_k j(t-k) \right)^2 \quad (38)$$

The solution of the minimization above is obtained by solving the normal equations

$$R_j(k) - \sum_{i=1}^p a_i R_j(i-k) = 0 \quad k = 1, \dots, p \quad (39)$$

where

$$R_j(k) = \frac{1}{N-k} \sum_{t=0}^{N-k-1} j(t)j(t+k)$$

and N is the number of data points in the analysis window. The estimate of G is given by

$$G = R_0 - \sum_{k=1}^p a_k R_k \quad (40)$$

We notice that the statistics of $j(t)$ needed for estimating its parameters are quadratic functions of $j(t)$.

Equations (36), (37), and (39), (40) give us the estimate of the parameters had we observed the complete data. Unfortunately, we only observe $y(t)$. However, using the *EM* idea, we can estimate in each step the necessary statistics of the complete data, given the observations and the previous value of the parameters.

Consider for simplicity the strict periodic model for the periodic part of $h(t)$. Given a current value of its parameters, $T^{(n)}$ and $a(t; T^{(n)})$, the periodic components $x(t)$ is completely defined, and will be denoted $x(t; T^{(n)}, a^{(n)})$. The *E* step of the algorithm will estimate the following statistics of $h(t)$:

$$\begin{aligned} E \{ h(t) | y(t); T^{(n)}, a(t; T^{(n)}), \phi^{(n)} \} &= x(t; T^{(n)}, a^{(n)}) \\ &+ E \{ e(t) | y(t); T^{(n)}, a(t; T^{(n)}), \phi^{(n)} \}. \end{aligned} \quad (41)$$

Similarly,

$$\begin{aligned} E \{ h(t)h(\sigma) | y(t); T^{(n)}, a(t; T^{(n)}), \phi^{(n)} \} &= x(t; T^{(n)}, a^{(n)}) \\ &+ x(\sigma; T^{(n)}, a^{(n)}) \\ &+ E \{ e(t) | y(t); T^{(n)}, a(t; T^{(n)}), \phi^{(n)} \} x(\sigma; T^{(n)}, a^{(n)}) \\ &+ x(t; T^{(n)}, a^{(n)}) E \{ e(\sigma) | y(t); T^{(n)}, a(t; T^{(n)}), \phi^{(n)} \} \\ &+ E \{ e(t)e(\sigma) | y(t); T^{(n)}, a(t; T^{(n)}), \phi^{(n)} \}. \end{aligned} \quad (42)$$

The necessary statistics of $j(t)$ to be estimated in the E step are

$$E\{j(t)|y(t); T^{(n)}, a(t; T^{(n)}), \phi^{(n)}\} \quad (43)$$

and

$$E\{j(t)j(\sigma)|y(t); T^{(n)}, a(t; T^{(n)}), \phi^{(n)}\}. \quad (44)$$

We observe that given $x(t; T^{(n)}, a^{(n)})$, we can subtract it from $y(t)$ to generate a signal $z(t)$. Given the parameters, the conditional expectation given $y(t)$ is equivalent to the conditional expectation given $z(t)$. We also observe that given the parameters, $z(t)$ is the sum of two zero-mean stochastic Gaussian processes, $e(t)$ and $j(t)$, whose second moments are given. Thus we can easily find the conditional expectation of $e(t)$ and $j(t)$ and also the conditional expectations of their quadratic terms.

Specifically, let σ_e^2 be the current variance of $e(t)$ and let $R_j(\phi^{(n)})$ be the current $N \times N$ covariance matrix of $j(t)$. Then the vectors $\hat{e} = [\hat{e}(0), \dots, \hat{e}(N-1)]^T$ and $\hat{j} = [\hat{j}(0), \dots, \hat{j}(N-1)]^T$ where $\hat{e}(t) = E\{e(t)|y(t); T^{(n)}, a(t; T^{(n)}), \phi^{(n)}\}$ and $\hat{j}(t) = E\{j(t)|y(t); T^{(n)}, a(t; T^{(n)}), \phi^{(n)}\}$ are given by

$$\hat{e} = [R_j(\phi^{(n)}) + \sigma_e^2 I]^{-1} \sigma_e^2 I \cdot z \quad (45)$$

$$\hat{j} = [R_j(\phi^{(n)}) + \sigma_e^2 I]^{-1} R_j(\phi^{(n)}) \cdot z \quad (46)$$

where $z = [z(0), \dots, z(N-1)]^T$.

The $N \times N$ matrices $\widehat{\bar{e}\bar{e}^T}$ and $\widehat{\bar{j}\bar{j}^T}$, whose t, s element is $E\{(e(t) - \hat{e}(t))(e(s) - \hat{e}(s))\}$ and $E\{(j(t) - \hat{j}(t))(j(s) - \hat{j}(s))\}$ are given by

$$\begin{aligned} \widehat{\bar{e}\bar{e}^T} &= \sigma_e^2 I - \sigma_e^2 I [R_j(\phi^{(n)}) + \sigma_e^2 I]^{-1} \sigma_e^2 I \\ &= \sigma_e^2 [R_j(\phi^{(n)}) + \sigma_e^2 I]^{-1} R_j(\phi^{(n)}) \end{aligned} \quad (47)$$

$$\begin{aligned} \widehat{\bar{j}\bar{j}^T} &= R_j(\phi^{(n)}) - R_j(\phi^{(n)}) [R_j(\phi^{(n)}) + \sigma_e^2 I]^{-1} R_j(\phi^{(n)}) \\ &= R_j(\phi^{(n)}) [R_j(\phi^{(n)}) + \sigma_e^2 I]^{-1} \sigma_e^2. \end{aligned} \quad (48)$$

The components of the estimated matrices above together with the components of the estimated vectors \hat{e} and \hat{j} and the current periodic signal estimate $x(t; T^{(n)}, a^{(n)})$ provide the conditional expectations of the necessary quadratic terms.

The conditional expectations above can be simplified if we can assume that the observation window length N is large enough, so we can work in the frequency domain, assuming that the Fourier coefficients of the signals are uncorrelated. The Fourier coefficients of, say, $z(t)$ are given by

$$Z(\omega) = \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} z(t) e^{-j\omega t}. \quad (49)$$

The conditional expectations of $E(\omega)$ and $J(\omega)$, the Fourier coefficient of $e(t)$ and $j(t)$, are

$$\hat{E}(\omega) = \frac{\sigma_e^2}{S_j(\omega; \phi^{(n)}) + \sigma_e^2} Z(\omega) \quad (50)$$

$$\hat{J}(\omega) = \frac{S_j(\omega; \phi^{(n)})}{S_j(\omega; \phi^{(n)}) + \sigma_e^2} Z(\omega) \quad (51)$$

where we recognize these expressions as the (noncausal) Wiener filter operations. For the necessary quadratic terms we have to calculate the conditional expectations $E\{\|E(\omega) - \hat{E}(\omega)\|^2\}$ $E\{\|J(\omega) - \hat{J}(\omega)\|^2\}$, given by

$$\begin{aligned} E\{\|E(\omega) - \hat{E}(\omega)\|^2\} &= \frac{\sigma_e^2 \cdot S_j(\omega; \phi^{(n)})}{\sigma_e^2 + S_j(\omega; \phi^{(n)})} \\ &= E\{\|J(\omega) - \hat{J}(\omega)\|^2\}. \end{aligned} \quad (52)$$

We now have all the conditional expectations needed for updating the parameters in the M step. The appropriate conditional expectations are substituted in the appropriate place in (36), (37), and (39), (40), instead of the unobserved statistics. These equations are then used to provide a new value for the parameters. The resulting algorithm is summarized and described explicitly below:

- 1) Start the algorithm by initial guess of the period $h(t)$ and the spectral parameters of $j(t)$. A plausible initial guess for T is the first peak of the autocorrelation of the observed signal. An averaging procedure like (10) can provide an initial estimate of the period waveform, with that T . Subtracting this waveform from the observed signal and solving the least squares prediction error equations on the residual will provide an initial estimate for the AR coefficients of $j(t)$.
- 2) Iterate as follows. At each iteration n an estimate of the parameters that determine the (almost) periodic signal (e.g., $T^{(n)}, \alpha^{(n)}, \dots$), an estimate of the periodic component $x^{(n)}(t) = x(t; T^{(n)}, \alpha^{(n)}, \dots, a^{(n)})$ and an estimate of the spectral parameters $\phi^{(n)}$ of $j(t)$ are all given. Thus, the E step will be
 - a) Subtract the periodic component, i.e., generate a signal $z(t)$

$$z(t) = y(t) - x^{(n)}(t). \quad (53)$$
 - b) Perform a Wiener filtering operation on $z(t)$ to generate an estimate of $j(t)$, $e(t)$, and their autocorrelation. Specifically, use the procedures described in (45)–(48), or (50)–(52).
 - c) Add the estimate of $e(t)$ to $x^{(n)}(t)$ to get an estimate of $h(t)$, and use the autocorrelation of $e(t)$ with $x^{(n)}(t)$ to get the estimate of the autocorrelation of $h(t)$.

Note that an estimate of the wide-band signal $j(t)$ is provided as a by-product of this E step.

The M step will be:

- a) Substitute the E step estimate of $h(t)$ in, e.g., (10) to get an estimate of the periodic waveform $a(t)$.
- b) Substitute the E step autocorrelation estimate of $h(t)$ in, e.g., (36) to get the updated estimate of T , and the other parameters that determine the (almost) periodic signal behavior.

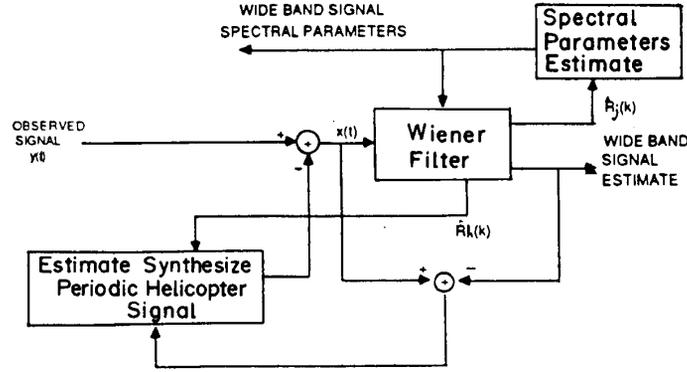


Fig. 3. The iterative algorithm for separation and parameter estimation.

- c) Use the E step estimate of the correlation of $j(t)$ in the normal equations to get an updated estimate of the AR parameters of $j(t)$.
- 3) End up, as a result, with new values for the (almost) periodic signal parameters, $T^{(n+1)}$, $\alpha^{(n+1)}$, \dots , the spectral parameters, $\phi^{(n+1)}$, of $j(t)$, and most importantly a new estimate of the periodic component $x^{(n+1)}(t)$ which together with the estimate of $e(t)$ from the E step provides an estimate of the signal $h(t)$.

This algorithm for separation and parameter estimation is summarized in Fig. 3.

B. Modification: Two Fundamental Frequencies

As mentioned above, the problem considered in this paper was motivated by the real situation where a helicopter acoustic signal is observed with additive wide-band jet plane signal. However, in many cases, the almost periodic helicopter signal contains two fundamental frequencies as a result of unsynchronized main and tail rotors. In this case the helicopter signal may be modeled as

$$h(t) = x_1(t) + x_2(t) + e(t) \quad (54)$$

where $x_i(t)$ is an almost periodic signal with parameters T_i , \dots and period waveform $a_i(t)$.

For simplicity, assume again that $e(t)$ is Gaussian and white. In this case if we observe $h(t)$, the estimate of the parameters associated with the two different fundamental frequencies is achieved by

$$\min_{T_1, \dots, a_1(t), T_2, \dots, a_2(t)} \int_{\mathcal{J}} [h(t) - x_1(t; T_1, \dots, a_1(t)) - x_2(t; T_2, \dots, a_2(t))]^2 \quad (55)$$

This minimization can be complicated and may require a multidimensional search, even when $h(t)$ is given. Thus we suggest the following iterative alternate minimization algorithm to solve (55): At each iteration a current estimate of the parameters of $x_1(t)$ is given. This estimate define a current value $x_1^{(n)}(t)$ for $x_1(t)$. Subtract this esti-

mate from $h(t)$ to get $h_1^{(n)}(t) = h(t) - x_1^{(n)}(t)$. Then, solve the least squares problem

$$\min_{T_2, \dots, a_2(t)} \int_{\mathcal{J}} (h_1^{(n)}(t) - x_2(t; T_2, \dots, a_2(t)))^2 \quad (56)$$

The solution of this problem is the current estimate of the parameters of $x_2(t)$, and can be used to define its current value $x_2^{(n)}(t)$. Subtract this estimate from $h(t)$ to get $h_2^{(n)}(t) = h(t) - x_2^{(n)}(t)$, and then solve the least square problem

$$\min_{T_1, \dots, a_1(t)} \int_{\mathcal{J}} (h_2^{(n)}(t) - x_1(t; T_1, \dots, a_1(t)))^2 \quad (57)$$

The solution of this problem provides the updated estimate of the parameters of $x_1(t)$ and defines its updated value $x_1^{(n+1)}(t)$. The iterations continue, reducing in each step the least square goal function of (55) until convergence.

This alternate minimization algorithm is easy to implement. Each of the least squares problems of (56), (57) is analogous to minimizing (20) which requires a search only for the parameters of a signal with one period.

When the helicopter signal is observed with additive wide-band signal, $h(t)$ is not available. However, we are using the EM algorithm. Given the complete data, i.e., given $h(t)$, we have to use the above alternate minimization algorithm. This will make the M step more complicated, but it still uses quadratic statistics of $h(t)$ in each iteration of the alternative minimization. Thus, the E step derived above, can be used without modification in this more general case.

IV. EXPERIMENTAL RESULTS

This study has been motivated by the real problem of separating the acoustic helicopter signal and the additive wide-band jet plane signal. Thus, it was important to test our algorithm using real helicopter signals and to find out that, indeed, it provides a valid answer to the studied problem.

Two helicopter test signals have been examined. The first, was a CH-47 helicopter signal whose rotors are synchronized and so it has a single fundamental frequency.

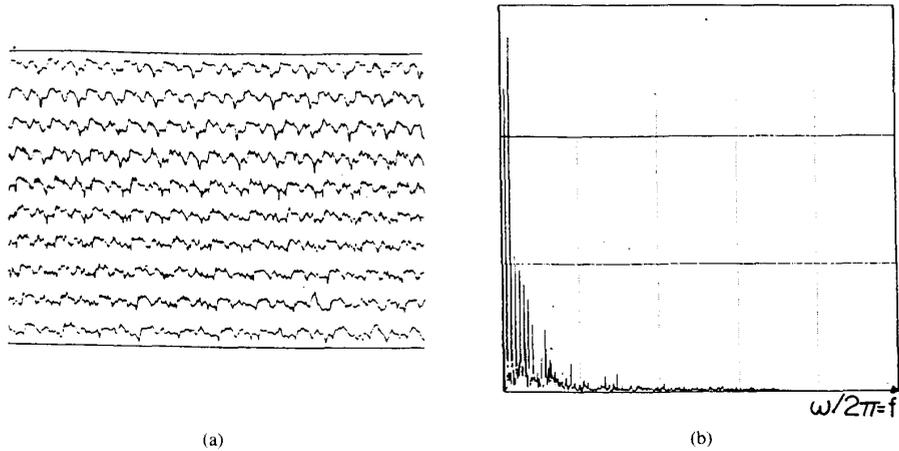


Fig. 4. CH-47: (a) Sample signal, (b) sample spectrum.

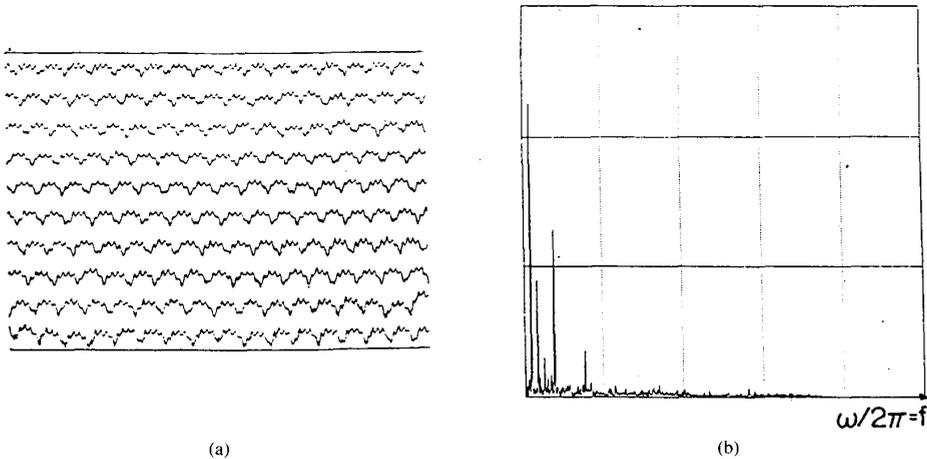


Fig. 5. UH-60: (a) Sample signal, (b) sample spectrum.

The second, was a UH-60 helicopter, in which two fundamental frequencies exist. Sample signals for both helicopters and the magnitude of their Fourier transform are shown in Figs. 4, 5.

As a first simple experiment the period of an observed CH-47 signal, assuming that it is periodic with a small error signal. We have calculated the least squares goal function suggested in (20) for various observation windows. The result is shown in Fig. 6. As one can see the specific signal we have used has a stable period, of approximately 92 ms, (fundamental frequency of 10.9 Hz). However, this does not mean that it is strictly periodic since the period waveform was varying along the test signal observations, due to fading (gain change) and a small Doppler shift. The estimate of the periodic signal itself is shown in Fig. 7. As we can see, most of the periodic components have been extracted successfully.

We have performed a similar experiment using the UH-60 signal. Here we had to consider the fact that this signal has two fundamental frequencies. Thus, we need the modification suggested for this case and iterated, can-

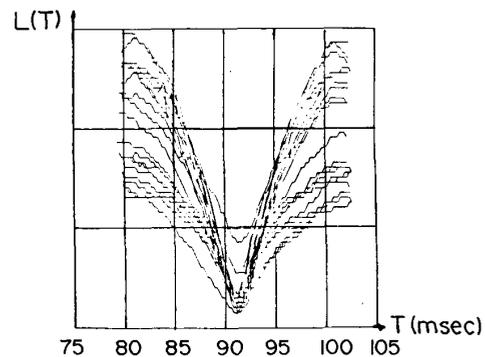


Fig. 6. CH-47: The goal function (square error) as a function of T . Different observation window lengths.

celing each periodic signal and calculating the goal functions $L(T_1)$ and $L(T_2)$ for each period estimate. These goal functions are shown in Fig. 8 for a few successive iterations. Again we clearly see in this example the period estimates. (The "uncertainty" in estimating T_2 reflected by

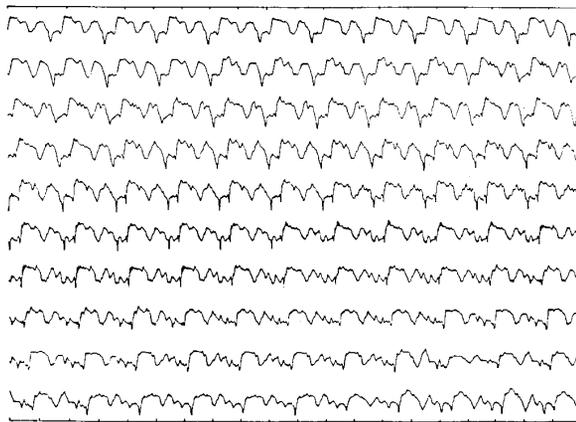


Fig. 7. Extracted CH-47 signal.

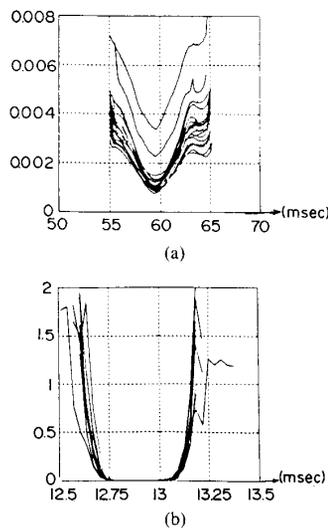


Fig. 8. UH-60: The goal function (square error) as a function of periods (a) T_1 and (b) T_2 : several iterations.

the flat shape of the goal function is not so severe if we normalize it by the higher fundamental frequency associated with this period.)

Next we have added a wide-band noise to the CH-47 helicopter, by generating sample functions of an AR process with various SNR. The AR signal level was controlled by the input variance G . The AR model order was 2. The parameters were chosen to provide a wide-band spectrum. We have implemented the algorithm described in Fig. 3. A total of five EM iterations were performed, where most of the improvement in estimation and extraction has been achieved in the first three iterations. The mean-square error of the estimated helicopter signal normalized by the energy of the true helicopter signal, which is the noise-to-signal ratio of the algorithm output, is shown as a function of the input SNR, in Fig. 9. The separated helicopter signal is shown, compared to the true helicopter signal in Fig. 10, for 0-dB SNR.

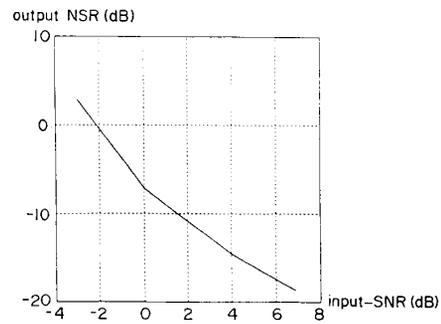


Fig. 9. CH-47 signal with an AR signal: Mean-square extraction error (NSR) as a function of input SNR.

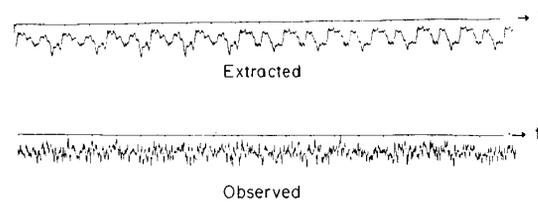


Fig. 10. CH-47 with AR: Extracted CH-47 signal at 0 dB.

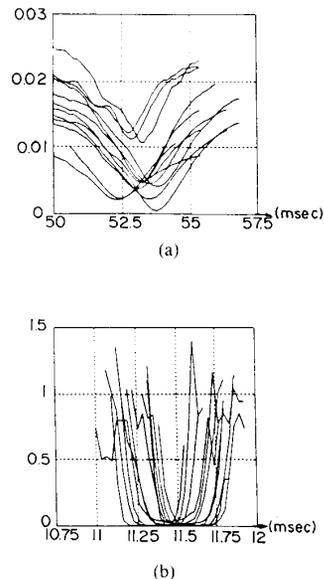


Fig. 11. UH-60 with a jet plane interference: Goal function (square error) as a function of periods (a) T_1 and (b) T_2 .

The most general experiment we performed was for the case of UH-60 helicopter in the presence of a jet plane signal. Not only did we have two fundamental frequencies, but the helicopter had a nonconstant radial velocity which led to nonstable periods. Since the periods were changing slowly, we used the methods described in (24), (25). Δ was on the order of 10 periods. The behavior of

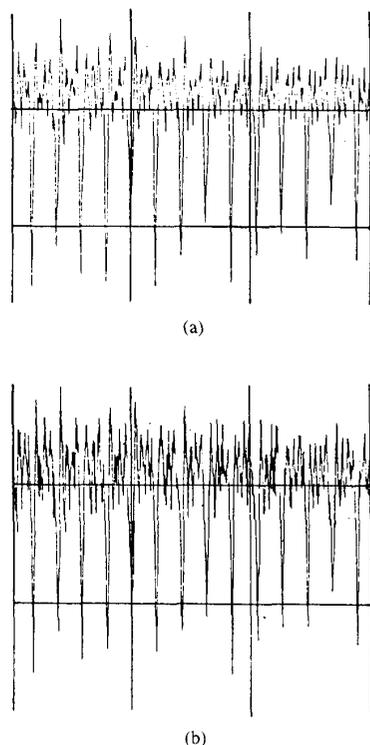


Fig. 12. UH-60 with real jet plane signal: (a) Observed and (b) extracted UH-60 signal.

the period goal functions $L(T_1)$ and $L(T_2)$ are shown in Fig. 11. The reconstructed helicopter signal is shown in Fig. 12.

The experiments reported here are by no means exhaustive. However, we may conclude, especially from the last example, performed using real signals, that the suggested algorithms are, indeed, capable to enhance the helicopter signals and to separate periodic and wide-band signals in a variety of cases. Additional experiments with these algorithms and a more comprehensive study of their performance is necessary and is left for further research.

APPENDIX A WEIGHTED PERIODIC SIGNAL

In this Appendix we present a method for estimating the parameters and the waveform of a signal that satisfies

$$x(t) = w_k \cdot x(t + kT) \quad (58)$$

i.e., each of its periods has a different gain. Unlike the discussion in Section II, the method suggested here also estimates the gains w_k .

Suppose first that the period time T is given. If the gains $\{w_k\}$ are known, the estimate of $a(t)$ is given by (13). On

the other hand, when the waveform $a(t)$ is given, the gain of the k th period can be estimated by

$$w_k = \frac{\int_{kT}^{(k+1)T} a(t) y(t) dt}{\int_{kT}^{(k+1)T} a^2(t) dt} \quad (59)$$

where, as before, (59) minimizes the least squares criterion

$$\int (y(t) - w_k \cdot a(t))^2 dt.$$

Equation (13) and (59) can define an iterative algorithm: We start with a guess of w_k , say all 1's. At each step we use the previous gain to find new $a(t)$ using (13) and we substitute the new waveforms in (59) to get w_k . We may restrict the gains to lie in an interval $1 - \alpha \leq w_k \leq 1 + \alpha$ by modifying the output of (59) as

$$w_k^{(\text{mod})} = \begin{cases} 1 - \alpha & \text{if } w_k < 1 - \alpha \\ w_k & \text{if } 1 - \alpha \leq w_k \leq 1 + \alpha \\ 1 + \alpha & \text{if } w_k > 1 + \alpha. \end{cases} \quad (60)$$

Now, if the period is not given one can find the best $\{w_k(T)\}$ and $a(t; T)$ for each T and, as before, search for the minimum of, say, a least square goal function

$$L(T) = \int \left(y(t) - \sum_k w_k(T) a(t; T) \right)^2 dt. \quad (61)$$

APPENDIX B

THE LIKELIHOOD OF THE OBSERVATIONS IN SECTION III

We have claimed that the direct ML estimate of both the parameters of the periodic signal and the parameters of the wide-band signal is complicated. This was the motivation for using the EM algorithm. To substantiate this claim we derive here an expression for the likelihood of the observations.

For simplicity we will assume that the observation window is long enough so we can work in the frequency domain, assuming that the different Fourier coefficients are uncorrelated. Thus our observations in the frequency domain are

$$Y(\omega_l) = H(\omega_l) + J(\omega_l) \quad (62)$$

where

$$Y(\omega_l) = \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} y(t) e^{-j\omega_l t} \quad \omega_l = l \frac{2\pi}{N} = l \frac{2\pi}{N}. \quad (63)$$

$J(\omega_l)$ is a zero-mean complex Gaussian random variable with variance $S_j(\omega_l; \phi)$. $H(\omega_l)$ is also Gaussian; its

mean is $X(\omega_j; T, \dots, a(t))$ where $X(\omega_j; T, \dots, a(t))$ is the Fourier transform of the periodic component $x(t; T, \dots, a(t))$, and its variance is σ_E^2 due to the white error signal $e(t)$. Thus $Y(\omega_j)$ is a Gaussian random variable with mean $X(\omega_j; T, \dots, a(t))$ and variance $\sigma_E^2 + S_j(\omega_j)$. Since all the frequency components are uncorrelated, the likelihood is proportional to

$$-\sum_l \log [S_j(\omega_j; \phi) + \sigma_E^2] + \frac{\|Y(\omega_j) - X(\omega_j; T, \dots, a(t))\|^2}{S_j(\omega_j; \phi) + \sigma_E^2}. \quad (64)$$

For ML estimation we have to maximize this expression with respect to ϕ and $T, \dots, a(t)$. Maximizing this expression can be complicated, and anyway, no explicit analytical solution is available.

APPENDIX C

THE EM ALGORITHM FOR MAXIMUM LIKELIHOOD ESTIMATION

The EM algorithm for maximum likelihood estimation is briefly summarized in this Appendix.

We denote by Y the data vector with the associated probability density $f_Y(y; \theta)$ indexed by the parameter vector θ where the possible parameter values are contained in a set Θ . Given an observed y , the ML estimate $\hat{\theta}_{ML}$ is the value of θ that maximizes the log-likelihood, that is

$$\hat{\theta}_{ML} = \arg \max_{\theta \in \Theta} \log f_Y(y; \theta) \quad (65)$$

Suppose that the data vector Y can be viewed as being incomplete, and we can specify some "complete" data X related to Y by

$$H(X) = Y \quad (66)$$

where $H(\cdot)$ is a noninvertible (many to one) transformation.

The EM algorithm is directed at finding the solution to (65); however, it does so by making an essential use of the complete data specification. The algorithm is basically an iterative method. It starts with an initial guess $\theta^{(0)}$, and $\theta^{(n+1)}$ is defined inductively by

$$\theta^{(n+1)} = \arg \max_{\theta \in \Theta} E \{ \log f_X(x; \theta) / y; \theta^{(n)} \} \quad (67)$$

where $f_X(x; \theta)$ is the probability density of X , and $E \{ \cdot / y; \theta^{(n)} \}$ denotes the conditional expectation given y , computed using the parameter value $\theta^{(n)}$. The intuitive idea is that we would like to choose θ that maximizes $\log f_X(x; \theta)$, the log-likelihood of the complete data. However, since $\log f_X(x; \theta)$ is not available to us (because the complete data is not available), we maximize instead its ex-

pectation, given the observed data y . Since we used the current estimate $\theta^{(n)}$ rather than the actual value of θ which is unknown, the conditional expectation is not exact. Thus the algorithm iterates, using each new parameter estimate to improve the conditional expectation on the next iteration cycle (the E step) and then uses this conditional estimate to improve the next parameter estimate (the M step).

The EM algorithm was presented in its general form by Dempster *et al.*¹ in [10]. The algorithm was suggested before, for specific applications, by several authors, e.g., [15]–[17]. The rate of convergence of the algorithm is linear [10], depending on the fraction of the covariance of the complete data that can be predicted using the observed data. If that fraction is small, the rate of convergence tends to be slow, in which case one could use standard numerical methods to accelerate the algorithm.

We note that the EM algorithm is not uniquely defined. The transformation $H(\cdot)$ relating X to Y can be any noninvertible transformation. Obviously, there are many possible "complete" data specifications that will generate the observed data. Thus, the EM algorithm can be implemented in many possible ways. The way $H(\cdot)$ is specified (i.e., the choice of the "complete" data) may critically affect the complexity and the rate of convergence of the algorithm.

REFERENCES

- [1] B. Gold and L. R. Rabiner, "Parallel processing techniques for estimating pitch period of speech in the time domain," *J. Acoust. Soc. Amer.*, vol. 46, Aug. 1969.
- [2] J. D. Wise, J. R. Caprio, and T. W. Parks, "Maximum likelihood speech estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-24, pp. 418–423, 1976.
- [3] L. R. Rabiner, "On the use of autocorrelation analysis for pitch detection," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-25, pp. 24–33, 1977.
- [4] J. S. Lim, A. V. Oppenheim, and L. D. Braida, "Evaluation of an adaptive comb filtering method for enhancing speech degraded by white noise addition," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-26, pp. 354–358, 1978.
- [5] M. V. Lowson, *Helicopter noise: Analysis, prediction, and method of reduction*, 1973.
- [6] R. F. Dwyer, "Extraction of helicopter-radiated noise by frequency domain processing," *J. Acoust. Soc. Amer.*, vol. 78, no. 1, pp. 95–99, 1985.
- [7] D. C. Rife and R. R. Boorstyn, "Multiple tone parameter estimation from discrete-time observations," *Bell Syst. Tech. J.*, vol. 55, no. 9, pp. 1389–1410, 1976.
- [8] A. Nehorai and B. Porat, "Adaptive comb filtering for harmonic signal enhancement," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1124–1138, 1986.
- [9] Meir Feder, "Iterative algorithms for parameter estimation with applications to signal processing," Ph.D. dissertation, Massachusetts Institute of Technology, 1987.
- [10] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc.*, ser. 39, pp. 1–38, 1977.

¹In [10] it is shown that each iteration increases the likelihood. However, there is an error in the convergence proof [10, theorem 2], pointed out by Wu [14]. The proper conditions that guarantee the convergence of the algorithm to a stationary point of the likelihood are given in [14].

- [11] S. O. Rice, "Mathematical analysis of random noise," *Bell Syst. Tech. J.*, vol. 14, pp. 465-471, 1947.
- [12] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York; McGraw-Hill, 1965.
- [13] R. H. Frazier, S. Samsam, L. D. Braidia, and A. V. Oppenheim, "Enhancement of speech by adaptive filtering," in *Proc. Int. Conf. Acoust., Speech, Signal Processing* (Philadelphia, PA), 1976.
- [14] C. F. J. Wu, "On the convergence properties of the EM algorithm," *Ann. Stat.*, vol. 11, pp. 95-103, 1983.
- [15] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains," *Ann. Math. Stat.*, vol. 41, pp. 164-171, 1970.
- [16] H. O. Hartley and R. R. Hocking, "The analysis of incomplete data," *Biometrics*, vol. 27, pp. 783-808, 1971.
- [17] T. Orchard and M. A. Woodsbury, "A missing information principle: Theory and applications," in *Proc. 6th Berkeley Symp. Math. Stat. Prob.*, vol. 1, pp. 697-715, 1972.



Meir Feder (S'81-M'87) received the B.Sc. and M.Sc. degrees (*summa cum laude*) from Tel-Aviv University, Tel-Aviv, Israel, and the Sc.D. degree from the Massachusetts Institute of Technology (M.I.T.), Cambridge, and the Woods Hole Oceanographic Institution (WHOI), Woods Hole, MA, all in electrical engineering in 1980, 1984, and 1987, respectively.

During 1987-1988 he was a Research Associate in the Department of Electrical Engineering and Computer Science at M.I.T. He then worked in Elbit Computers, Haifa, Israel. In October 1989 he joined the faculty of the Department of Electrical Engineering—Systems, Tel-Aviv University. He was also a Visiting and Guest Investigator at WHOI during the summers of 1983 and 1988-1991. His research interests include data and signal compression, topics in information theory, signal and image processing, and sonar processing.