JOINT DETECTION AND HIGH RESOLUTION ML ESTIMATION OF MULTIPLE SINUSOIDS IN NOISE

M. D. Macleod
(mdm@eng.cam.ac.uk) Signal Processing Group, Department of Engineering, University of Cambridge, UK.

ABSTRACT

Harmonic analysis, the analysis of signals which consist of a sum of sinusoids (or complex sinusoids) with additive white or colored noise, is a much studied problem, with many important applications. Nevertheless, existing approaches have significant limitations. In many, the model order (number of sinusoids) is assumed known, and in most cases Additive White Gaussian Noise (AWGN) is assumed. We present a method for jointly determining the model order and estimating the sinusoid parameters in white or colored noise. It uses the notch periodogram in an iterative detection and estimation algorithm. It uses an explicit detection test based on an estimate of the noise PDS, which is obtained by smoothing the logarithm of the notch periodogram.

1. INTRODUCTION

Let $\mathbf{x} = \left[x(0), x(1), \ldots, x(N\text{-}1)\right]^T$ be the signal consisting of M complex sinusoids (cisoids) in noise:

$$\mathbf{x} = \sum_{m=1}^{M} a_m \mathbf{e}(\omega_m) + \mathbf{v} = \mathbf{E}(\boldsymbol{\Omega})\mathbf{a} + \mathbf{v} \tag{1.1}$$

where a_m and ω_m are the complex amplitude and normalized frequency of the m^{th} cisoid,

$$\mathbf{e}(\omega_m) = \left[e^0, e^{j\omega_m}, e^{j2\omega_m}, \ldots, e^{j(N-1)\omega_m}\right]^T,$$

$\mathbf{v} = \left[v(0), v(1), \ldots, v(N\text{-}1)\right]^T$ is the noise vector,

$$\mathbf{E}(\boldsymbol{\Omega}) = \left[\mathbf{e}(\omega_1), \mathbf{e}(\omega_2), \ldots, \mathbf{e}(\omega_M)\right]^T,$$

$$\boldsymbol{\Omega} = [\omega_1, \omega_2, \ldots, \omega_M]^T \text{ and } \mathbf{a} = [a_0, a_1, \ldots, a_{N-1}]^T.$$

The problem considered in this paper is to jointly estimate the Power Density Spectrum of the noise (which may be colored), determine M, and estimate the signal parameter vectors $\mathbf{a}$ and $\boldsymbol{\Omega}$. If $\mathbf{v}$ is white and Gaussian and M is known, the Maximum Likelihood Estimates of $\mathbf{a}$ and $\boldsymbol{\Omega}$ are found [1] by solving the least squares minimization

$$\min_{\boldsymbol{\Omega},\mathbf{a}} J(\boldsymbol{\Omega}, \mathbf{a}) \equiv \left\|\mathbf{x} - \mathbf{E}(\boldsymbol{\Omega})\mathbf{a}\right\|^2. \tag{1.2}$$

For given $\boldsymbol{\Omega}$, the optimum $\mathbf{a}$ is given [1] by

$$\mathbf{a} = \left(\mathbf{E}(\boldsymbol{\Omega})^H \mathbf{E}(\boldsymbol{\Omega})\right)^{-1} \mathbf{E}(\boldsymbol{\Omega})^H \mathbf{x}. \tag{1.3}$$

Hence substituting (1.3) into (1.2) yields an equivalent minimization of an objective function of $\boldsymbol{\Omega}$ only,

$$\min_{\boldsymbol{\Omega}} J_1(\boldsymbol{\Omega}) \equiv \left\|\mathbf{x} - \mathbf{P}_{\mathbf{E}(\boldsymbol{\Omega})}\mathbf{x}\right\|^2 = \left\|\mathbf{P}^{\perp}_{\mathbf{E}(\boldsymbol{\Omega})}\mathbf{x}\right\|^2, \tag{1.4}$$

where $\mathbf{P}_{\mathbf{E}(\boldsymbol{\Omega})} = \mathbf{E}(\boldsymbol{\Omega})\left(\mathbf{E}(\boldsymbol{\Omega})^H \mathbf{E}(\boldsymbol{\Omega})\right)^{-1} \mathbf{E}(\boldsymbol{\Omega})^H$ is the projector onto $\operatorname{span}\left(\mathbf{E}(\boldsymbol{\Omega})\right)$ and $\mathbf{P}^{\perp}_{\mathbf{E}(\boldsymbol{\Omega})} = \mathbf{I} - \mathbf{P}_{\mathbf{E}(\boldsymbol{\Omega})}$ is the orthogonal projector. An alternative approach is to form a minimization problem using a prediction error matrix, and solve it iteratively using the Steiglitz-McBride algorithm (1974), KiSS algorithm (Kumaresan, Scharf, Shaw, 1986), IQML algorithm (Bresler, Macovski, 1986), or (in a variant form) the IFA (Kay, 1984). However the algorithm in this paper, like that in [2,3], minimizes (1.4) directly in terms of $\boldsymbol{\Omega}$ and uses a different iterative approach, with important benefits.

2. THE NOTCH PERIODOGRAM

Assume that the true values of the first $M-1$ frequencies are known, and form the vector $\boldsymbol{\Omega}_V = [\omega_1, \ldots, \omega_{M-1}]^T$. To determine the one remaining frequency ω, we can use the result [2] that $J_1(\boldsymbol{\Omega}) = J_1(\boldsymbol{\Omega}_V) - P_V(\omega; \boldsymbol{\Omega}_V)$, where

$$P_V(\omega; \boldsymbol{\Omega}_V) = \frac{\left|\mathbf{e}^H(\omega)\mathbf{P}^{\perp}_{\mathbf{E}(\boldsymbol{\Omega}_V)}\mathbf{x}\right|^2}{\left\|\mathbf{P}^{\perp}_{\mathbf{E}(\boldsymbol{\Omega}_V)}\mathbf{e}(\omega)\right\|^2} \tag{2.1}$$

is called the notch periodogram of the data vector $\mathbf{x}$ with respect to the notch set $\boldsymbol{\Omega}_V$. Hence the optimum value of ω is that which maximizes $P_V(\omega; \boldsymbol{\Omega}_V)$. This leads to the idea [2,3] of using the notch periodogram iteratively.

In [3] Hwang and Chen describe a combined detection & estimation algorithm, which has two phases. The initialization phase starts by identifying the frequency ω_{01} of the peak in the standard periodogram. Next, with $\boldsymbol{\Omega}_1 = \omega_{01}$ as the notch set, the new peak frequency ω_{02} of the notch periodogram is found. Then with $\mathbf{\Omega}_2 = [\omega_{01}, \omega_{02}]^T$ as the notch set, the frequency ω_{03} of the peak of the notch periodogram is determined, and so on. At each iteration, the model order M is increased only if it results in a reduction in the following information-theoretic criterion [4]

$$\mathrm{EDC}(k) = NJ_1(\mathbf{\Omega}_k) + \frac{3k}{2}\alpha \ln N \quad (2.2)$$

in which the first term is the log-likelihood (ignoring a constant term) and the second a penalty term; otherwise the initialization phase is terminated. In [3], the value $\alpha = 2$ in (1.6) is empirically found to be optimum. Interestingly, a more recent Bayesian model order criterion for this problem [5] corresponds to $\alpha = 5/3$ in (2.2).

The second phase of the algorithm in [3] refines the frequency estimates using the Alternating Notch Periodogram Algorithm (ANPA) [2]. Each frequency in turn is removed from the notch set, the peak of the resulting notch periodogram is found, and its frequency replaces the removed frequency in the notch set. This optimization process (which is of the type known as univariate search or 'cyclic descent' [6]) is repeated until convergence.

A separate application of the notch periodogram is a detection test by Li and Djurić [7] for closely spaced sinusoids in white noise. First assume one tone, plus noise $\mathbf{v}$ of variance σ^2. Assume the frequency ω_1 of the single sinusoid is correctly estimated and, with notch frequency ω_1, the frequency of the peak of the notch periodogram in the range $(\omega_1 - \pi/N) < \omega < (\omega_1 + \pi/N)$, i.e. one "bin" width centered on ω_1, is ω_P. Then $2P_1(\omega_P;\omega_1)/\sigma^2$ is [7] approximately distributed (in the tails) as χ_2^2. Thus a threshold γ_1 can be defined so that if $P_1(\omega_P;\omega_1) < \gamma_1$ we conclude there is only one sinusoid; otherwise we conclude that there is at least one more sinusoid close in frequency. [7] then describes an extension of this idea to test for more than two sinusoids. A new notch set is formed either from the frequencies of the two largest peaks in $P_1(\omega;\omega_1)$ in the above frequency range or, if there is only one peak at ω_P, as $\mathbf{\Omega}_2 = [\omega_1, \omega_P]^T$. The maximum of the new notch periodogram in the same frequency range is again tested and if it is less than γ_1 we conclude there are only two sinusoids. Otherwise the number of notches is increased in the same way and the test repeated until the test fails. Where σ^2 is not known, [7] proposes estimating it by averaging the notch periodogram over the frequency "bins" not close to the notch frequencies.

This detection test is part of an overall framework [7] in which each distinct peak in the original periodogram is processed in the way described above, but while the k^th peak is being analyzed, notches are also placed at the other peaks of the original periodogram, and the noise estimation procedure is modified to exclude those peaks.

3. DEVELOPMENT OF THE NEW METHOD

We have previously described fast frequency-domain approaches to the harmonic estimation problem [8,9]. These algorithms were based on subtraction of the estimated tones, rather than their removal by orthogonal projection as in the notch periodogram. However, we found it difficult to devise statistically satisfactory detection tests for that approach, for the high resolution case. We therefore decided to combine the detection test [7] with an iterative estimator based on the notch periodogram, as in [3]. (A further reason for this choice were the limitations in the performance of information-theoretic stopping criteria such as (2.2); see [7] and also our findings below.) A key benefit is that we can now devise an extension to the detection test to accommodate colored noise. If the "noise floor" $S_N(\omega)$,

$$S_N(\omega) = \frac{1}{N} E\left\{\left|\mathbf{e}^H(\omega)\mathbf{v}\right|^2\right\}, \quad (3.1)$$

where E signifies Expectation, varies only "slowly" with frequency, then over a small frequency range it is approximately constant. Hence if it can be estimated, we may use $S_N(\omega)$ instead of σ^2 in the detection test.

3.1 Investigation of existing algorithms

The detection test [7] is satisfactory for deciding between 0, 1, and more-than-1 tone. However it has limitations. The first may be illustrated using the (often used [1]) test signal used in [7], that is $y(n) =$

$$\exp(j2\pi 0.5n) + b\exp(j2\pi 0.51n + \phi) + v(n), \quad (3.2)$$

where $n = 0,...,24$ and $\phi = \pi/4$. Also define Signal to Noise Ratio (as in [1, 7]) as $10\log_{10}(1/\mathrm{var}(v(n)))$ dB. The first notch is placed at the frequency ω_1 of the single peak (near the mean of the two true frequencies), and the second notch is placed very close to the first. Thus the two notches are not close to the two true frequencies. At a SNR of 20 dB, the maximum SNR used in [7], the resulting residual error is not sufficiently large to cause erroneous detection of a third tone. However, at higher SNR (e.g. 30dB) the test erroneously detects 3 or 4 tones.

Next, note the signal (3.2) has a "best case" phase relationship. The worst case phase [9] in (3.2) is $\phi = -\pi/4$

rather than $+\pi/4$, and it makes both detection and estimation more difficult (for any algorithm) - a higher SNR is needed to detect two tones. After the first notch is placed (again near the mean of the two true frequencies), the residual notch periodogram has smaller total energy. And, although the residual has two peaks, they are well outside the frequency range of the test [7], further reducing its detection power. Even if that frequency range is increased, the test suffers because the power in the residual is spread in frequency. We found that better detection performance results from the idea [12] of using a lowered initial threshold then, if that threshold is exceeded, adding a new notch, optimizing the notch frequencies, and then retesting each tone by removing each notch in turn.

Also, [7] often places new notches at the same frequency as existing notches, or very close, and this causes numerical problems in the notch periodogram algorithm. Suitable minor algorithm modifications can prevent this.

We then investigated the combined detection-estimation algorithm [3], which works well in many circumstances. However in difficult cases two problems arise. First, the detection test based on the EDC (2.2) has lower detection power than that in [7]. For example, if the signal (3.2) is used with $\phi = -\pi/4$ at SNR 40dB, only one tone is detected by the EDC method, whereas two are detected by [7]. This may be in part because the peak frequencies found in the initialization phase of [3] are biased due to the presence of other tones nearby in frequency. However, the same bias is found for $\phi = \pi/4$, but in that case two tones are successfully detected.

Second, for high resolution estimation, the convergence of the ANPA algorithm is fast for "good" phase relationships, such as (3.2) with $\phi = \pi/4$, which was one of the test cases used in [3]. However, like all univariate optimization methods, it converges much more slowly when the Hessian matrix of the error surface has a high eigenvalue ratio, as happens for example when $\phi = -\pi/4$ is used in (3.2). For example, at 40dB SNR, with $\phi = \pi/4$ ANPA converges completely within 5 iterations; but with $\phi = -\pi/4$, (and forced detection of two tones) convergence is slower, taking over 100 iterations.

3.2 A new noise floor estimation algorithm

If the noise floor (3.1) is assumed to be white, then its mean value σ^2 may be estimated [7] by averaging the notch periodogram, excluding frequencies close to detected tones. During initial detection, the resulting estimated is biased high by the currently undetected tones. This may in certain cases reduce detection performance. One way to reduce this bias is to estimate σ^2 using the mean of the logarithm of the notch periodogram:

$$\hat{\sigma}^2 = 1.776 \exp\left(N^{-1} \sum_{k=0}^{N-1} \log P_V\left(k\omega_0 ; \mathbf{\Omega}_V\right) \right) \tag{3.3}$$

where $\omega_0 = 2\pi/N$ (or, if a zero-padding factor of Z is used, $\omega_0 = 2\pi/(ZN)$). Although this estimator has a higher variance than the (unrealizable) sample-mean of the periodogram of noise alone, it has a far lower bias from any undetected tones. Thus the probability of failed detection of new peaks is reduced. We may extend this approach to estimate a colored noise floor (3.1), by forming a weighted local average in the frequency domain, instead of a global average:

$$\hat{S}_N\left(m\omega_0\right) = 1.776 \exp\left(\sum_{k=-K}^{K} w_k \log P_V\left(\left(m+k\right)\omega_0 ; \mathbf{\Omega}_V\right) \right)$$

where $w_k, k = -K, \ldots, K$ is a unit sum window function (such as a scaled Hanning window). Such an approach has been proposed [10] as a smoothed estimator of a broadband spectrum, but not for its ability to estimate a noise floor spectrum in the presence of interfering tones.

4. THE OVERALL ALGORITHM

Our overall algorithm involves iterative noise floor estimation, using the methods described above, detection of new tones, using an enhanced detection test, and refinement of the notch frequencies of the notch periodogram by non-linear optimization. A crucial issue, which significantly affects performance, is the sequence in which detection, frequency estimation and notch frequency optimization are carried out.

To overcome the detection test problems described in section 3.1: (a) we avoid to the need [7] to start by identifying "l distinct peaks" in the periodogram, which requires a separate detection test. Instead, as each new peak is detected, we either assign it to an existing "cluster" if it is within 2 "bins" (i.e. $4\pi/N$ rad/samp) of any notch frequency in an existing cluster, or form a new single-tone cluster. This change also removes the (rigid) one bin frequency range in [7], which (as we noted) reduces detection power in certain cases; (b) we re-optimize the notch frequencies after each new tone is detected; and (c) we use a lower initial threshold, but then retest with a higher threshold after notch frequency optimization.

Because the ANPA algorithm has very slow convergence under some conditions, we use a BFGS Quasi-Newton algorithm (Matlab routine fminunc) for the optimization. If the number of tones is large, the computational load of this optimization is high. We therefore investigated methods of reducing it. Because "clusters" are separated from each other in frequency they have low cross-interference [9], so a single optimization of all notches may be replaced by sequential optimization of each cluster in turn independently with little loss of performance; this works well and significantly reduces computation load.

The following sequence was found to be the best:
1. compute notch periodogram, noise floor and threshold
2. compare new peak in notch periodogram with threshold; if peak below threshold, terminate.
3. assign new peak frequency to existing cluster (if within 2 bins of any notch in cluster) or form new cluster.
4. optimise notches in each cluster in turn (fminunc).
5. go to 1.

4.1 Joint detection and estimation results

Example 1: two tones, equation (3.2) using *worst case* phase $\phi = -\pi/4$, white noise, SNR=40dB. In 100 trials, exactly 2 tones were detected in all cases. The mean estimated frequencies and rms errors were:

True freq (bins)	12.5	12.75
Mean result (bins)	12.499	12.761
rms error (bins)	0.0374	0.0452

Example 2: blocklength N=64, four tones, using worst case phase relationships [9] between *all* tones, white noise. In 50 trials, exactly 4 tones were detected in all cases. The mean estimated frequencies and rms errors were:

Tone SNR (dB)	24	33	33	38
True freq (bins)	7.516	7.816	15.016	15.216
Mean result	7.519	7.815	14.974	15.234
rms error	0.015	0.006	0.107	0.083

4.2 Computation issues

We used fast frequency domain algorithms which we have developed for computing the notch periodogram [11]. It is necessary to limit the minimum separation of notches to avoid numerical problems due to ill-conditioning in the computation of $\mathbf{P}_{\mathbf{E}(\mathbf{\Omega})}$ and (2.1).

For example 1 (high resolution, two tones, N=24) each trial required between 100 and 210 notch periodogram evaluations, and took approximately 4 seconds on a Dell notebook CpiAD400XTB (Pentium II processor).

For example 2 (high resolution, 4 tones, N=64) each run required typically 200 and 400 notch periodogram evaluations, and computation took approximately 20 seconds (mean).

Note that although the second-order optimization method (fminunc) is more complex than simple ANPA univariate optimization, the number of notch periodogram evaluations (which dominates computation time) is actually reduced for these worst case phase examples.

5. CONCLUSION

Our algorithm incorporates: the concept of iterative optimization of the notch periodogram from [3], but with an enhanced optimization strategy; an enhanced detection test based on the notch periodogram; estimation of a (possibly colored) noise floor (when required) through the smoothed log periodogram [10]; and fast frequency-domain notch periodogram algorithms [11]. It performs well, even for challenging test cases.

REFERENCES

1. S. M. Kay, *Modern Spectral Estimation*, Prentice-Hall, Englewood Cliffs, NJ, 1988.
2. J.-K. Hwang & Y.-C. Chen, "Superresolution Frequency Estimation by Alternating Notch Periodogram", *IEEE Trans. SP*, **41**, no. 2, Sept 1993, pp.727-741.
3. J.-K. Hwang & Y.-C. Chen, "A Combined Detection-Estimation Algorithm for the Harmonic Retrieval Problem", *Signal Processing*, **30**, no.2, 1993, pp.177-197.
4. L.C.Zhao, P.R.Krishnaiah & Z.D.Bai, "On detection of the number of signals in presence of white noise", *J. Multivariate Anal.*, **20**, 1986, pp.1-25.
5. P. M. Djurić, "A Bayesian model order determination rule of harmonic signals", Proc. ISCAS-1995, pp. 2273-6.
6. P. Bloomfield, *Fourier Analysis of Time Series: An Introduction*, Wiley, New York, 1976.
7. H.T. Li, P.M. Djurić, "A novel approach to detection of closely spaced sinusoids", *Signal Processing*, **51**, (2) pp. 93-104, Jun 1996.
8. M. D. Macleod, "Fast Nearly-ML Estimation of the Parameters of Real or Complex Single Tones or Resolved Multiple Tones", *IEEE Trans SP*, 1998, **46**, 1, pp.141-148.
9. M. D. Macleod, "High Resolution Nearly-ML estimation of Sinusoids in Noise using a Fast Frequency Domain Approach", *Proc EUSIPCO 98*, (Sept 98), pp. 1849-1852.
10. G. Wahba, "Automatic Smoothing of the Log Periodogram", *J Amer Statist Assoc*, **75**, 1980, pp. 122-132.
11. M. D. Macleod, "A Fast Frequency Domain Notch Periodogram Algorithm", *accepted for publication in Signal Process.*
12. I. J. Clarke, Personal Communication.