Audio Engineering Society
Convention Paper
Presented at the 111th Convention
2001 September 21–24 New York, NY, USA

Bayesian Harmonic Analysis for Audio Testing and Measurement

Thomas J. Loredo
Muse
Ithaca, NY 14850, USA
loredo@museweb.com

ABSTRACT
Many common audio test and measurement procedures require characterization of the output signal of the device under test in terms of harmonic (sinusoidal) components and residual noise when the device processes sinusoidal input signals. This work uses the Bayesian approach to statistical inference to address such problems as parameter estimation problems when discrete samples of the output signal are given. In the resulting Bayesian harmonic analysis the power spectrum computed from the discrete-time Fourier transform appears as the logarithm of the posterior probability for the frequency of a single sinusoid rather than as an estimate of the signal spectrum; more complicated functions of the transform arise when analyzing signals with multiple sinusoids. Problems such as spectral leakage are addressed by nonlinear processing of the Fourier transform, offering several advantages over methods that use (linear) windowing of data.

INTRODUCTION

Many common audio test and measurement procedures require characterization of the output signal of the device under test in terms of harmonic (sinusoidal) components and residual noise when the device processes a harmonic input signal. Examples include measurement of frequency and phase response, signal-to-noise ratio, total harmonic distortion, and intermodulation distortion; see [1, 2, 3] for excellent overviews of standard methods for performing such measurements. This work uses the Bayesian approach to statistical inference to address such problems as parameter estimation problems when discrete samples of the output signal are given. Jaynes first adopted this approach in a 1987 analysis of acoustic chirp from bats [4]; Bretthorst has applied it with considerable success to analysis of decaying sinusoidal signals in nuclear magnetic resonance (NMR) data from chemical analyses of materials [5, 6]. In Jaynes's analysis, the power spectrum or periodogram derived from the discrete-time Fourier transform

(DTFT) appears, not as an estimate of the signal spectrum, but as the logarithm of the posterior probability density for the frequency of a sinusoid in the sampled data.

This paper explores the consequences of a similar analysis of data from tone-based testing of audio systems. The resulting algorithms process the DTFT in a nonlinear manner, in contrast to traditional methods that smooth the power spectrum via data tapering or windowing. The nonlinear processing produces accurate estimates of signal characteristics (e.g., amplitudes of fundamentals and harmonics and measurement of the noise level) without requiring smoothing to reduce "spectral leakage," and thus without the loss of resolution resulting from smoothing. When the signal has more than one sinusoidal component (e.g., for distortion measurements), Bayesian harmonic analysis identifies nonlinear combinations of the real and imaginary parts of the DTFT at multiple frequencies that can produce accurate estimates of signal characteristics even when the frequencies lie within a single power spectrum peak, provided the SNR is high enough. When the frequencies are well-separated, the multiple-component analysis effectively reduces to the simpler single-component analysis.

This brief report motivates the Bayesian approach, outlines the general algorithm, and presents a few sample calculations analyzing simulated data to demonstrate the approach. Detailed derivations and development of additional algorithms will appear in a future publication.

AUDIO TESTING AS STATISTICAL INFERENCE

To motivate what follows, we first recall some basic facts about sampled time series. Suppose one has uniformly spaced samples of the amplitude $h(t)$ of a signal that is a continuous function of time, t. Denote the sample values by $d_i = h(t_i)$, with $t_i = i\Delta t$, with Δt the sampling interval and $i = 0$ to $N-1$. Suppose further that the signal is known to be bandlimited to the Nyquist interval, and that it is known to be comprised of a number M of discrete sinusoids of unknown frequencies, phases, and amplitudes. If the samples are noiseless and $M = 1$ (a single sinusoid is present), then with just $N = 3$ samples we can perfectly infer the frequency, amplitude, and phase of the sinusoid because only the correct sinusoid will pass exactly through all the samples. This is true for any frequency from DC to the Nyquist frequency. If $M > 1$, we need more samples ($3M$), but it remains possible to infer the frequencies, phases, and amplitudes of the sinusoids perfectly with relatively few samples. The method of least squares can be used to find the sinusoid parameters, with no uncertainty in the estimated parameters.

When analyzing data from real systems, even if we believe there is only a single sinusoid in the data we must use many samples because noise complicates our inferences. The data do not directly give us $h(t_i)$; rather,

$$d_i = h(t_i) + n_i, \quad (1)$$

where n_i is the (unknown) noise contribution to datum i. As a result, a candidate sinusoid that passes close to but not exactly through the d_i cannot be ruled out as a possible $h(t)$; the difference between the candidate sinusoid and the d_i values may be due to noise. The measurement problem has qualitatively changed from one of simple logic to one of *statistical inference*. One must develop estimators for the sinusoidal parameters that acknowledge the presence of noise and account for the resulting uncertainty in inferences of the signal parameters.

Although many basic discussions of audio testing methodology make little if any reference to statistics, the mathematical methods of audio testing have statistical origins. In particular, techniques such as power spectrum estimation and data windowing have their origins in the conventional frequentist or sampling theory approach to statistics ("frequentist" here refers to the interpretation of probability as giving the long-run frequency of occurrence of an outcome in repeated trials). But there is another approach to statistical inference—the Bayesian approach—that leads to different procedures in many problems. Historically, it is the original approach to statistical inference, dating back to the work of Laplace, Bayes, Gauss and their contemporaries in the late 1700s and early 1800s. This approach fell out of favor in the late 1800s, setting the stage for the development of the frequentist approach by Fisher, Pearson, Neymann and others in the early 1900s. But in recent decades there has been a resurgence of interest in the Bayesian approach, due in large part to the rapid growth of computing power making feasible the sometimes complicated calculations it requires.

In this work we explore the implications of the Bayesian approach for audio testing, viewing testing as a statistical inference problem (parameter estimation). As background, we first review how the frequentist and Bayesian approaches differ.

In frequentist statistics, one chooses a default "null" hypothesis, H_0 (e.g., a model specifying the signal $h(t)$), and selects a statistic $S(D)$ (a function of the data, D) that measures departure of the data from the predictions of H_0. The value of S found using the observed data, $S(D_{\rm obs})$, is used to make inferences (e.g., estimate or constrain a parameter, decide whether to reject H_0). Using the sampling distribution for the data, $p(D|H_0)$, one then calculates the distribution for the statistic given the null, $p(S|H_0)$ (analytically or via Monte Carlo simulation, generating hypothetical data sets and evaluating $S(D)$ for each simulated data set). Sums and moments of this distribution provide measures of the long-run performance of basing inferences on $S(D_{\rm obs})$. These include the bias of an estimator, the confidence level associated with a confidence region, or the significance level ("false alarm" probability) of a hypothesis test.

In the Bayesian approach, one must specify at least two competing hypotheses for the data (e.g., a single parameterized model with the parameters indexing the hypotheses; or two or more competing models that may or may not have unknown parameters). One then uses the rules of probability theory to calculate the posterior probabilities for the hypotheses given the observed data, $p(H_i|D_{\rm obs}, I)$, with the hypotheses denoted by H_i, and the background information and assumptions used in the calculation denoted by I (including, e.g., distributional assumptions for the noise). According to Bayes's theorem,

$$p(H_i|D_{\rm obs}, I) \propto p(H_i|I)p(D_{\rm obs}|H_i, I), \quad (2)$$

where $p(H_i|I)$ is the prior probability for H_i, and $p(D_{\rm obs}|H_i, I)$, considered as a function of H_i, is the likelihood for H_i. For final inferences, one calculates sums and moments of (2). For example, to report probabilities for hypotheses, one must sum the right hand side to find the normalization constant. Another common example arises when estimating parameters in a model. Typically, the model, M, has both interesting parameters, s (e.g., signal parameters), and uninteresting parameters, b (e.g., background DC level or possibly the noise level). The hypotheses needed to predict the data are indexed jointly by s and b, but ultimately we are directly interested only in s. Inferences for s alone can be found by integrating the posterior distribution for s and b over b, or *marginalizing* over b:

$$p(s|D_{\rm obs}, M) = \int db\, p(s, b|D_{\rm obs}, M); \tag{3}$$

$p(s|D_{\rm obs}, M)$ is called the marginal distribution for s.

Two points of departure are apparent from this brief sketch of the two approaches (see [9, 10] for further discussion of the following points). The first concerns the choice of statistic. In the frequentist case, specifying a good statistic for a nontrivial problem is a difficult art. In the Bayesian approach, once a hypothesis space is specified, probability theory automatically identifies what functions of the data to use to discriminate between the hypotheses (i.e., the functions that appear in the likelihood). This automatic behavior comes at the cost of having to specify alternative hypotheses (some frequentist calculations can proceed without specifying an alternative to the null hypothesis, e.g., goodness-of-fit tests). Second, the two approaches use the sampling distribution for the data very differently. In frequentist calculations, the hypothesis is fixed, and sums and integrals are calculated in the sample space of hypothetical data. In Bayesian calculations, the data are fixed to the observed values, and sums and integrals are calculated in the hypothesis or parameter space. Consequently, even when the same statistics are used in both approaches, qualitatively different results can be found.

A BAYESIAN LOOK AT THE POWER SPECTRUM

As an illustration of how these differences manifest themselves in a simple setting, consider the problem of characterizing a periodic signal $h(t)$ using the uniformly sampled data d_i described above. For simplicity, we assume independent Gaussian distributions for the noise with zero mean and common standard deviation, σ (white noise).

To formulate a frequentist solution to this problem, we must choose a statistic. An informal rule that often guides the choice of statistic in a frequentist calculation might be summarized as follows: "Do to the data what you would do to the signal that produced it." Given a periodic function $h(t)$, an unknown period could be readily identified by examining the amplitude of the Fourier transform of $h(t)$. This suggests choosing a statistic based on the squared amplitude of the DTFT of the data,

$$I(f) = \frac{1}{N}\left[\sum_i d_i \cos(2\pi f t_i)\right]^2 + \frac{1}{N}\left[\sum_i d_i \sin(2\pi f t_i)\right]^2. \tag{4}$$

Viewed as a continuous function of the frequency f, we call $I(f)$ the (continuous) *periodogram*. In a frequentist calculation, we will be interested in how it behaves as the values of the N data vary through repeated observation. Since there are only N data, there must only be at most N "pieces of information" in the continuous function $I(f)$. Actually, there are $N/2+1$ (nearest integer if N is odd) values of $I(f)$ at equally spaced frequencies that determine the entire function. These values can be found using the discrete Fourier transform (DFT) of the data to calculate the power spectrum at $N/2+1$ Fourier frequencies, $I_N(f_j) \equiv I(f_j)$, where $f_j = 2\pi j/T$, with T equal to the duration spanned by the data and $j = 0$ to $N/2$. The Fourier power spectral density (PSD) that is conventionally used is $I_N(f_j)$, with a possible subtraction of an average term, and with various normalizations adopted (to simplify its statistical properties). For simplicity, we call $I_N(f_j)$ the PSD.

From the perspective of frequentist statistics, an important virtue of focusing on the Fourier frequencies is that the values of I_N at these frequencies are statistically independent under the null hypothesis of a constant (perhaps zero) signal (plus noise). But an unfortunate consequence of there only being $N/2+1$ Fourier frequencies is that the expected behavior of the PSD differs depending on whether the unknown true period of the signal lies exactly on or away from a Fourier frequency. This is illustrated in Figure 1. Figure 1a shows the PSD calculated from data with a signal at a Fourier frequency and a signal-to-noise of 10; Figure 1b shows the PSD calculated from similar data, but with the signal frequency midway between two Fourier frequencies. *Spectral leakage* is apparent; when the true frequency is not a Fourier frequency, power "leaks" to neighboring frequencies, reducing the amplitude of the PSD peak, and broadening it. This complicates the interpretation and use of the PSD for both detection and estimation. Conventional remedies for leakage use windowing or tapering of the data (essentially a linear averaging process) to reduce leakage at non-Fourier frequencies, at the expense of spreading the signal power when the signal is at or near a Fourier frequency. (More detailed discussion is available in [2, 3, 12, 13]; Priestly [13] gives a thorough discussion of the statistical underpinnings of frequentist periodogram methods.)

Now consider the problem from the Bayesian perspective. We cannot try to detect a periodic signal without specifying a model for one; obviously, there are many possible periodic models. As a starting point, presume $h(t)$ is a single sinusoid of unknown amplitude A, frequency f, and phase ϕ; then

$$d_i = A\cos(2\pi f t_i + \phi) + n_i. \tag{5}$$

This model, M_1, has three parameters. The likelihood function for the parameters, $\mathcal{L}(f, A, \phi) = p(D_{\rm obs}|f, A, \phi, M_1)$, is, assuming Gaussians for the error probabilities,

$$\mathcal{L}(f, A, \phi) \propto \exp\left[-\frac{1}{2}\chi^2(f, A, \phi)\right], \tag{6}$$

where

$$\chi^2(f, A, \phi) = \sum_i \frac{[d_i - A\cos(2\pi f t_i + \phi)]^2}{\sigma^2}. \tag{7}$$

To estimate the frequency of a periodic signal, we integrate the resulting posterior distribution over A and ϕ. Using flat priors for A and ϕ, Jaynes (1987) and Bretthorst (1989) found that for large N,

$$p(f|D_{\rm obs}, M_1) \propto \exp\left[\frac{I(f)}{\sigma^2}\right], \tag{8}$$

where the continuous periodogram, $I(f)$, arises automatically from the algebra *given the single sinusoid model*. If instead of estimating f we are seeking to detect a signal, we must compare M_1 to an alternative model, M_0, assuming, say, only a DC signal is present (with noise). Comparison of these models requires integration of the likelihoods associated with each over *all* parameters.

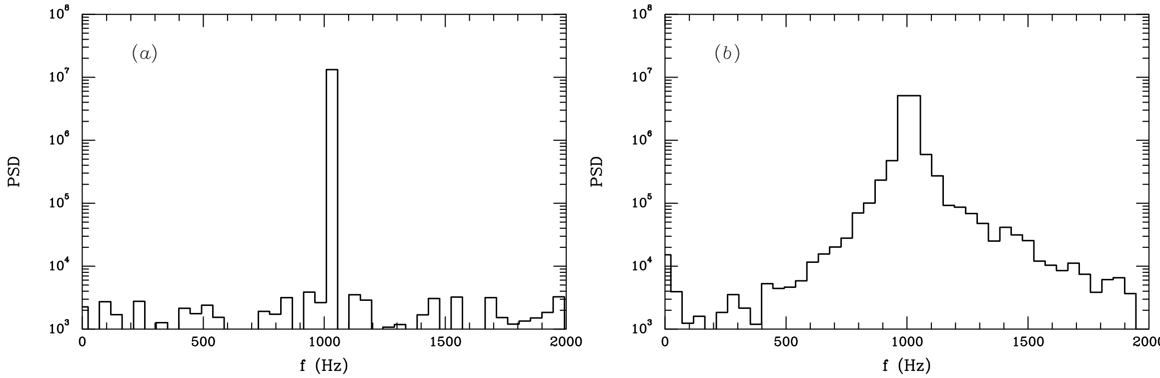

Fig. 1: Leakage in the PSD. (a) The PSD (up to 2 kHz) for data simulated with a weak sinusoidal signal and added Gaussian noise; $S/N = 5$, with 1024 samples at a 48 kHz sampling rate. The sinusoid frequency is at the Fourier frequency nearest 1 kHz (1031.25 Hz). (b) As in (a), but the frequency (1008 Hz) is between two Fourier frequencies.

Figure 2 illustrates some aspects of the Bayesian procedure. Figure 2a shows the continuous periodogram for the same data used to produce the solid curve in Figure 1a (signal at a Fourier frequency). Dots highlight the values at Fourier frequencies (the values plotted in Fig. 1a). Figure 2b shows a similar plot, corresponding to Figure 1b (signal at a non-Fourier frequency). Although the values at Fourier frequencies exhibit very different behavior in Figures 1a and 1b, the continuous periodograms are qualitatively very similar. The insets in the figures show the marginal posterior distributions for the frequency in each case, calculated using (8). These distributions are extremely sharp and narrow in both cases, and very accurately pinpoint the correct frequency. The sidelobes and other structure evident in the continuous periodograms are exponentially attenuated. Detection probabilities (for determining whether a periodic signal is present), found by integrating the exponentiated periodogram over f, similarly exhibit comparable performance for Fourier and non-Fourier frequencies.

We can now see important differences between Bayesian and frequentist use of Fourier methods. In Bayesian spectrum analysis, the periodogram arises automatically from consideration of a specific time-domain model—a single sinusoid. In frequentist statistics it is known that the PSD is closely related to the residuals found from least-squares fitting of a sinusoid (e.g., [14]), but the PSD is viewed in a more general fashion as a somewhat distorted estimate of the (continuous) power spectrum of the underlying signal. The sources of the distortion are the finite span of the data, and the discrete sampling; these cause spectral leakage and other problems.

From the Bayesian viewpoint, there is no spectral leakage problem associated with non-Fourier frequencies. The continuous periodogram has a complicated shape for *all* possible signal frequencies; the complications are merely hidden if one focuses only on the Fourier frequencies. This focus arises in the frequentist approach because of its reliance on sample space integrals; only periodogram values at Fourier frequencies are independent. Since the Bayesian calculation instead requires parameter space (e.g., frequency) integrals, the continuous periodogram is of interest, with Fourier frequencies playing no special role. Moreover, the complicated shape of the periodogram, though a consequence of the finite and discrete nature of the data, is not to be viewed as arising from convolution of an underlying spectrum with window and sampling functions. Rather, the shape conveys information about how the finite and discrete nature of the data can confuse one's inferences about a single sinusoid when noise is significant (in which case the sidelobes will not be as attenuated as in the examples above). The periodogram is interpreted as the logarithm of the marginal distribution for the frequency of a simple sinusoidal signal, not as an estimate of the PSD of the signal. When there is significant evidence for a sinusoid, the sidelobes and leakage in the periodogram are eliminated, not by linear smoothing (which decreases resolution), but by exponential attenuation (which sharpens the peak).

A further consequence of the Bayesian calculation is that the PSD is identified as the statistic of interest for a signal model with only a *single sinusoid*, calling into question the adequacy of the PSD for more complicated signals. For example, Bretthorst ([5]; § 6.3) shows that a model presuming two sinusoids leads to a different statistic. When the frequencies of the sinusoids are well separated, the statistic can be accurately estimated from the values of the periodogram at the two frequencies of interest (the frequencies produce distinct peaks); but when the frequencies are close, the periodogram fails. Instead one must calculate a nonlinear combination of the real and imaginary parts of the DTFT evaluated at the two frequencies, not just its squared modulus at a single frequency (as in the PSD). Using this statistic, the Bayesian procedure can estimate two frequencies accurately even when they lie well within a single periodogram peak (see below for an example). In other words, for signals more complicated than a single sinusoid, there is important information in the *phases* of the Fourier transform that can be extracted by time-domain Bayesian modeling.

Bretthorst [6, 7] has extended these calculations to develop methodology for analysis of NMR data that is now in wide use. NMR hardware produced by the Varian Corporation now ships with software implementing Bretthorst's methods; the success of these methods was recently reported in the science press [8]. The author has developed similar techniques in astronomy for time series sampled with Poisson rather than

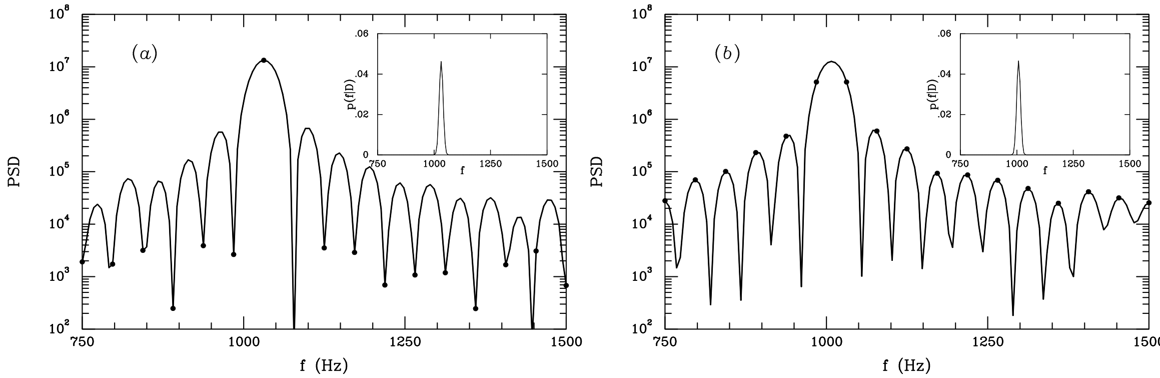

Fig. 2: Details of the continuous periodogram. (a) Periodogram near its peak, for the data used for Fig. 1(a) (at a Fourier frequency); dots show values of the (discrete) PSD. Inset shows the posterior distribution for the frequency found by exponentiating the periodogram scaled by the noise variance. (b) As in (a), but for the data used for Fig. 1(b) (signal frequency between Fourier frequencies).

Gaussian statistics (see, e.g., [11]), and is currently applying the Jaynes/Bretthorst methodology to the problem of detecting extrasolar planets using radial velocity and astrometric measurements of motions of nearby stars (this work is currently being prepared for publication). The present paper outlines application of the approach to selected audio test and measurement tasks that seek to decompose the signal from the device under test into a finite number of sinusoids. We call the resulting family of algorithms *Bayesian harmonic analysis* (BHA).

OUTLINE OF THE ALGORITHM

A detailed derivation of the general algorithm for Bayesian harmonic analysis will appear in a forthcoming publication. It is a special case of more general algorithms derived by Bretthorst [5]. For this report, we merely summarize the algorithm as a "recipe" whose elements should appear familiar to readers acquainted with Fourier analysis and linear least squares.

We model the signal as a sum of M sinusoids,

$$\begin{aligned} h(t) &= \sum_\alpha A_\alpha \cos(2\pi f_\alpha t - \phi_\alpha) \\ &= \sum_\alpha \left[C_\alpha \cos(2\pi f_\alpha t) + S_\alpha \sin(2\pi f_\alpha t)\right], \end{aligned} \quad (9)$$

where in the second line we replaced the original amplitude and phase with two amplitudes. This simplifies some of the later mathematics; the original model has two nonlinear parameters (f_α and ϕ_α) for each harmonic, but the reparameterized model has only one nonlinear parameter, the frequency. We will sometimes denote the frequencies collectively by $f = \{f_\alpha\}$, and similarly with the amplitudes A, S, and C. Note that we use greek indices to label the harmonics (going over 1 to M), reserving lower case roman indices for the data (going over 0 to $N-1$). Our goal is to develop algorithms for estimating the frequencies f and the amplitudes S and C (from which we can easily recover A and ϕ).

The sets of values of the cos and sin functions evaluated at the N sample times can be thought of as the components of vectors in an N-dimensional space. Denote these vectors by $\vec{c}_\alpha$ and $\vec{s}_\alpha$, with components $c_{\alpha i} = \cos(2\pi f_\alpha t_i)$, etc.. There are $2M$ such vectors. To simplify subsequent notation, denote them collectively by $\vec{g}_A$, where the A index runs from 1 to $2M$, and $\vec{g}_1 = \vec{c}_1$, $\vec{g}_2 = \vec{s}_1$, $\vec{g}_3 = \vec{c}_2$, etc.. Note that these vectors are functions of the frequencies f. Associated with these vectors are amplitudes B_A, with $B_1 = C_1$, $B_2 = S_1$, $B_3 = C_2$, etc..

The first ingredient we must calculate to implement BHA is the $2M \times 2M$ metric matrix η for the $2M$-dimensional subspace spanned by the model vectors. Its components are just the dot products of the model vectors,

$$\begin{aligned} \eta_{AB} &= \vec{g}_A \cdot \vec{g}_B \\ &= \sum_i g_{Ai} g_{Bi}. \end{aligned} \quad (10)$$

This is a symmetric, positive-definite matrix. Since the vectors are functions of f, the metric is as well. The sums in equation (10) are sums of products of sines and cosines over evenly spaced arguments, and can be calculated analytically.

The data samples can also be thought of as the components of a vector in the same N-dimensional space the models live in, with $\vec{d} = \{d_i\}$. The second ingredient we need is the collection of projections of the models on the data,

$$\begin{aligned} P_A &= \vec{d} \cdot \vec{g}_A \\ &= \sum_i d_i g_{Ai}. \end{aligned} \quad (11)$$

The projections are also functions of f. Since the components of the $\vec{g}$ vectors are just cosines and sines, the projections are just the real and (negative) imaginary parts of the DTFT of the data. If we evaluate them only at Fourier frequencies, they are given by the DTF of the data; but the Fourier frequencies will often form too crude a grid. We can efficiently interpolate between the Fourier frequencies by using a zero-padded DFT.

With these ingredients—the f-dependent metric and projections—we can now state the algorithm. For each f of interest, the estimated amplitudes are given by

$$B = \eta^{-1} P, \tag{12}$$

where η^{-1} is the matrix inverse of the metric, and B and P are $2M$-dimensional column vectors with components B_A and P_A. The components of η^{-1} are the covariances of the amplitude estimates; in particular, the square roots of the diagonal components give the individual uncertainties for each amplitude. One can combine these to find the uncertainty for the overall amplitude of each sinusoidal component.

To estimate the frequencies, calculate the *sufficient statistic*

$$S(f) = \sum_A B_A P_A. \tag{13}$$

Use this to calculate the squared residual of the fit to the data,

$$r^2(f) = \sum_i d_i^2 - S(f), \tag{14}$$

From this the marginal probability density for the frequencies is given by

$$p(f|D) = K \exp\left[-\frac{r^2(f)}{2\sigma^2}\right], \tag{15}$$

where K is a normalization constant that can be found by integrating the remaining factor over f. The most probable choice of f is the one that maximizes this probability, and thus the one that minimizes $r^2(f)$ or, equivalently, maximizes $S(f)$. Our uncertainty in f can be quantified by using the distribution in equation (15) to find a standard deviation for f or the size of a region that encloses, say, 95% of the probability.

For $M = 1$ (a single sinusoid), it is straightforward to show that $S(f_1)$ is proportional to the standard Fourier power spectrum when N is large (for large N the off-diagonal terms in the metric become negligible, leaving a simple diagonal matrix; see [5]). In this sense, the Bayesian approach derives the power spectrum. For $M = 2$, there are two frequencies to estimate, and the sufficient statistic is now a two-dimensional function, $S(f_1, f_2)$. One can show that when the two frequencies are well-separated (many Fourier frequencies apart), $S(f_1, f_2) \approx S(f_1) + S(f_2)$. Thus, the usual power spectrum has all the information needed to get the sufficient statistic. But when the frequencies are close, this decomposition fails, and one must use the full algorithm to make accurate inferences of the sinusoid parameters.

EXAMPLE CALCULATIONS

As a simple illustration of BHA in action, we consider here a simplified version of measurement of SMPTE intermodulation distortion (IMD). We simulated data containing the SMPTE test tones at $f_1 = 60$ Hz and $f_2 = 7$ kHz with a $4:1$ amplitude ratio; the simulated data also contained a single, very weak mock distortion product at $f_3 = f_2 - f_1 = 6940$ Hz. A real system would have other distortion products; here we limit ourselves to one only to simplify the plots and discussion. The amplitude of the distortion component was 3×10^{-5} of the amplitude of the 7 kHz component, representing IMD distortion at the 0.003% level (-90.46 dB). We simulated data sets with varying levels of noise. For all of the analyses reported here, $N = 1024$ data samples were used at a 48 kHz sampling rate. Note that this corresponds to less than 22 ms of sampling time, corresponding to just 1.28 periods of the 60 Hz modulation (a nonintegral number of cycles). The Fourier spacing is about 47 Hz, so the distortion product power is in a frequency bin immediately adjacent to the test signal's frequency bin in the FFT of the data.

Figure 3a shows the power spectrum of data simulated with a signal-to-noise ratio $S/N = 80$ dB (σ for the noise set to 10^{-4} times the amplitude of the 7 kHz test tone). The effects of spectral leakage are evident. There is no obvious indication of the presence of a distortion product; indeed it seems hopeless to expect to find such a product at -90 dB in these data.

Figure 3b shows contours of the BHA sufficient statistic as a function of the two upper frequencies with f_1 fixed at the known value of 60 Hz. The contours are chosen so that the innermost contour bounds a region with more than 99% of the marginal probability density for the frequencies. Despite the discouraging appearance of the power spectrum, it is evident that the two frequencies can be easily resolved by the Bayesian analysis, with frequency resolution orders of magnitude finer than the Fourier spacing for the 7 kHz signal. The inferred amplitude of the distortion product is 0.0031(5)%, with the digit in parentheses indicating the uncertainty in the last digit of the estimate. This agrees very well with the true value.

Figure 4 shows how the algorithm performs as a function of S/N. For each of a number of values of S/N, we simulated 20 data sets and used BHA to estimate the IMD level for each; the estimates are plotted as dots on the Figure. The algorithm produces useful estimates for S/N as low as 75 dB, which seems usefully low considering that the distortion product itself is below -90 dB and that very short and nonsynchronous data samples were used. It may appear that at lower S/N the algorithm misleads the user into thinking a larger distortion product is present than is actually present. But this is not so; for each estimate, the algorithm also provides an uncertainty (not shown on the plot). For the low S/N estimates, the estimated amplitude itself is typically smaller than one or two times its uncertainty, indicating that no definite distortion product was detected. This criterion can be made more precise within the Bayesian framework (using Bayesian model comparison to compare a two-frequency model to the three-frequency model assumed here). But in practice the informal rule of rejecting measurements smaller than a few times their uncertainty is usually adequate.

A future publication will elaborate on these developments, presenting a more complete derivation of Bayesian harmonic analysis and details of applications to several other common audio measurements.

The author gratefully acknowledges many conversations over the last decade with the late Ed Jaynes and especially with Larry Bretthorst that clarified his understanding of Bayesian inference with superposed nonlinear models (of which BHA is a special case).

REFERENCES

[1] R. C. Cabot, "Comparison of Nonlinear Distortion Measurement Methods," in *Proc. AES 11th Int. Conf. on Audio Test & Measurement,* pp. 53–65 (1992).

[2] R. C. Cabot, "Fundamentals of Modern Audio Measurement," *J. Audio Eng. Soc.* (Sep. 1999).

[3] B. Metzler, *Audio Measurement Handbook*, Beaverton, OR: Audio Precision (1993).

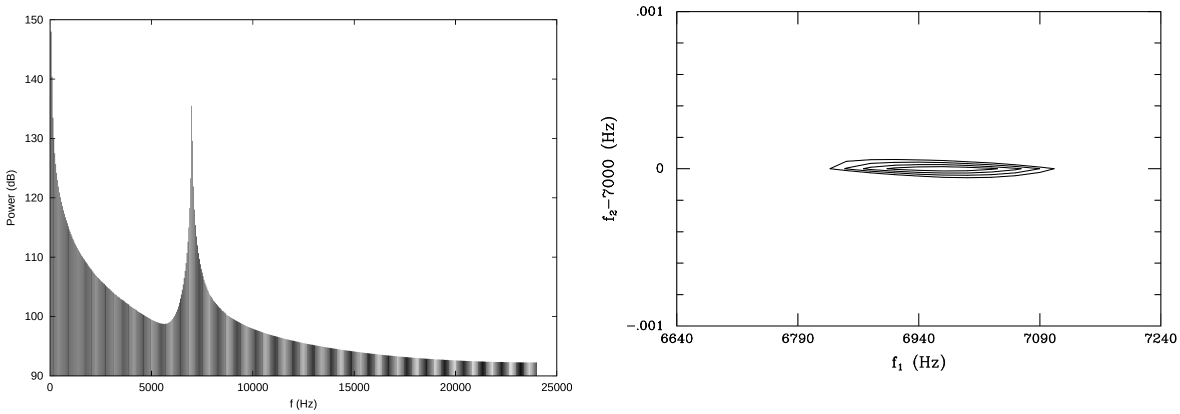

Fig. 3: (a, left) Periodogram of simulated data for the simplified SMPTE IMD measurement with $S/N = 80$ dB described in the text. Significant spectral leakage prevents measurement of the distortion product 90 dB below the 7 kHz test signal. (b, right) Contours of the Bayesian posterior probability density for the frequencies of the 7 kHz test signal and the distortion product at 6940 Hz, using the same data as used for (a).

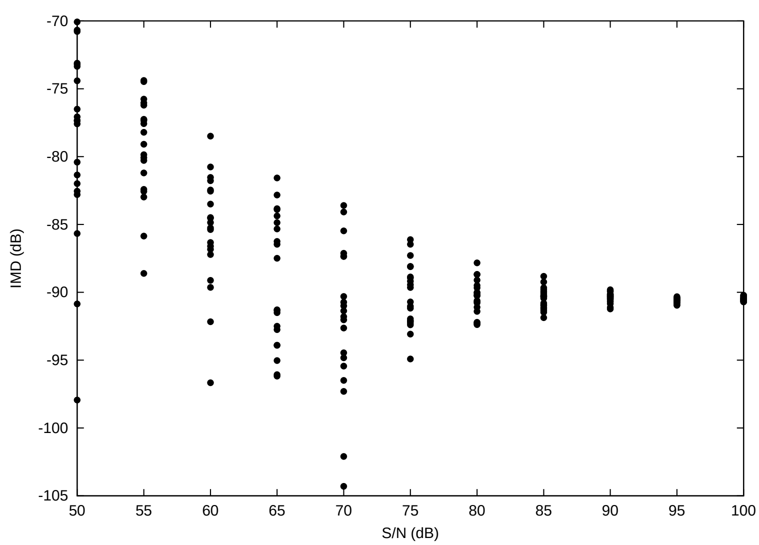

Fig. 4: Bayesian estimates of the amplitude of a single SMPTE IMD distortion component for data simulated with various S/N.

[4] E. T. Jaynes, "Bayesian Spectrum and Chirp Analysis," in *Maximum Entropy and Bayesian Spectral Analysis and Estimation Problems*, ed. C. R. Smith and G. J. Erickson, D. Reidel, Dordrecht-Holland, pp. 1–37 (1987).

[5] G. L. Bretthorst,, *Bayesian Spectrum Analysis and Parameter Estimation*, New York: Springer-Verlag (1988).

[6] G. L. Bretthorst, "Bayesian Analysis III: Applications to NMR Signal Detection, Model Selection and Parameter Estimation," *J. Magn. Reson.*, **88**, 571–595 (1990).

[7] G. L. Bretthorst, "Bayesian Analysis IV. Noise And Computing Time Considerations," *J. Mag. Res.*, **93**, pp. 369–394 (1991).

[8] D. Malakoff, "Bayes offers a 'New' Way to Make Sense of Numbers," *Science*, 286, 1460–1464 (1999).

[9] T. J. Loredo, "The Promise of Bayesian Inference for Astrophysics," in Statistical Challenges in Modern Astronomy, ed. E. Feigelson & G. Babu, NY:Springer, pp. 275–297 (1992).

[10] T. J. Loredo, "Computational Technology for Bayesian Inference," in *ADASS (Astronomical Data Analysis & Software Systems) VIII,* ed. R. Crutcher & D. Mehringer, San Francisco: Astronomical Society of the Pacific (1999).

[11] P. Gregory and T. J. Loredo, "A new method for detection of a periodic signal of unknown shape and period," *Astrophys. J.*, **398**, pp. 146–168 (1992).

[12] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge: Cambridge Univ. Press (1992).

[13] M. B. Priestly, *Spectral Analysis and Time Series*, London: Academic Press (1981).

[14] N. R. Lomb, "Least-squares frequency analysis of unequally spaced data," *Astrophys. & Space Sci.*, **39**, 447–462 (1976).