2000

Bayesian Spectrum Analysis for Laser Vibrometry Processing

Walter F. Buell
Institute for Advanced Physics

Bradley Allan Shadwick
University of Nebraska-Lincoln, shadwick@unl.edu

Robert W. Farley
Electronics and Photonics Laboratory The Aerospace Corporation

Bayesian Spectrum Analysis for Laser Vibrometry Processing

Walter F. Buell,[1,2*] B. A. Shadwick,[1]

and

Robert W. Farley[2]

[1] *Institute for Advanced Physics*
10875 U.S. Hwy. 285, Suite 199
Conifer, CO 80433

[2] *Electronics and Photonics Laboratory*
The Aerospace Corporation, M2-253, P.O. Box 92957
Los Angeles, CA 90009-2957

April 24, 2000

Revised September 22, 2000

ABSTRACT

Laser vibration sensing provides a sensitive non-contact means of measuring vibrations of objects. These measurements are used in industrial quality control and wear monitoring as well as the analysis of the vibrational characteristics of objects. In laser vibrometry, the surface motion is monitored by heterodyne laser Doppler velocimetry, and the received heterodyne signal is sampled to produce a time-series which is processed to obtain a vibrational spectrum of the object under test. Laser vibrometry data has been processed with a traditional FM discriminator approach and by spectrogram and time-frequency distribution processing techniques. The latter techniques have demonstrated improved performance over the FM discriminator method, but do not take full advantage of the prior knowledge one has about the signal of interest. We consider here a statistical signal processing approach to laser vibrometry data. In this approach the quantities of interest are the frequencies of vibration, while the phase and quadrature amplitudes are considered nuisance parameters. Because of the optimal use of prior knowledge about the laser vibrometry signal, the frequencies can be determined with much greater precision and greater noise immunity than using Fourier- or time-frequency-based approaches. Furthermore, the statistical approach is known to have superior performance when the data extends over a small number of vibrational periods. We illustrate the method with data from a fiber-optic laser Doppler velocimeter. Our results show that while the choice of processing method for determining the instantaneous velocity is relatively unimportant, the Bayesian method exhibits superior performance in determining the vibrational frequency.

I. INTRODUCTION

Laser vibrometry (LV) provides a sensitive non-contact means of measuring vibrations of objects. In a monostatic LV measurement, a laser beam illuminates an object of interest, which may be stationary or in motion, and the returned scattered light is mixed with a local oscillator derived from the same laser. The instantaneous beat frequency provides a measurement of the surface velocity which may be extracted from the time-series data by signal processing (for example a Fourier transform). The time evolution of the beat frequency then contains information about the state of motion (vibration or other time-varying velocity) of the target, which can be extracted by further processing. Traditional methods for processing LV data include the FM discriminator method[1], spectrogram processing (a Fourier method)[1] and time-frequency distributions[2]. Differences have been shown among the performance of several LV processing techniques, and it is natural to ask whether statistical signal processing approaches would further improve sensitivity or accuracy for laser vibration sensing. For example, Bayesian spectrum analysis has been demonstrated[3] to significantly improve resolution in processing nuclear magnetic resonance (NMR) data, where a time-series is processed to infer a molecular spectrum.

In order to be of utility for laser vibration sensing, a signal processing method must operate well in the presence of noise, be robust to speckle broadening and laser linewidth, and be computationally efficient. It should also yield useful spectra in a small number of vibrational periods, especially when the vibrational period is long or if the measurement time is limited. This latter requirement poses a significant constraint on fast Fourier transform (FFT) approaches because the frequency resolution is roughly equal to the inverse of the measurement time. Statistical spectrum estimators such as maximum entropy methods often perform much better than estimators based on the FFT in such situations.

In this paper we compare Bayesian signal processing to Fourier transform methods (in the spirit of spectrogram processing). The experimental data are derived from laser velocimetry of a rotating drum with a diffuse scattering surface. We find that, at least in the regime of high carrier-to-noise ratio, the FFT is sufficient to determine the instantaneous surface velocity. When determining the vibrational frequency, however, we find that the Bayesian frequency estimator dramatically outperforms the FFT approach. Section II describes our experimental arrangement. Section III provides an overview of the statistical processing methods we consider and describes our implementation. We present our results in Section IV, comparing the Bayesian method to FFT processing.

II. EXPERIMENT

To evaluate the performance of the statistical processing methods, we used data from a fiber-optic laser Doppler velocimeter developed for *in situ* fluid flow measurements (see Figure 1).[4] In the experiment, the target was a rotating wheel coated with a diffuse reflecting surface and driven with an optical chopper-wheel motor.

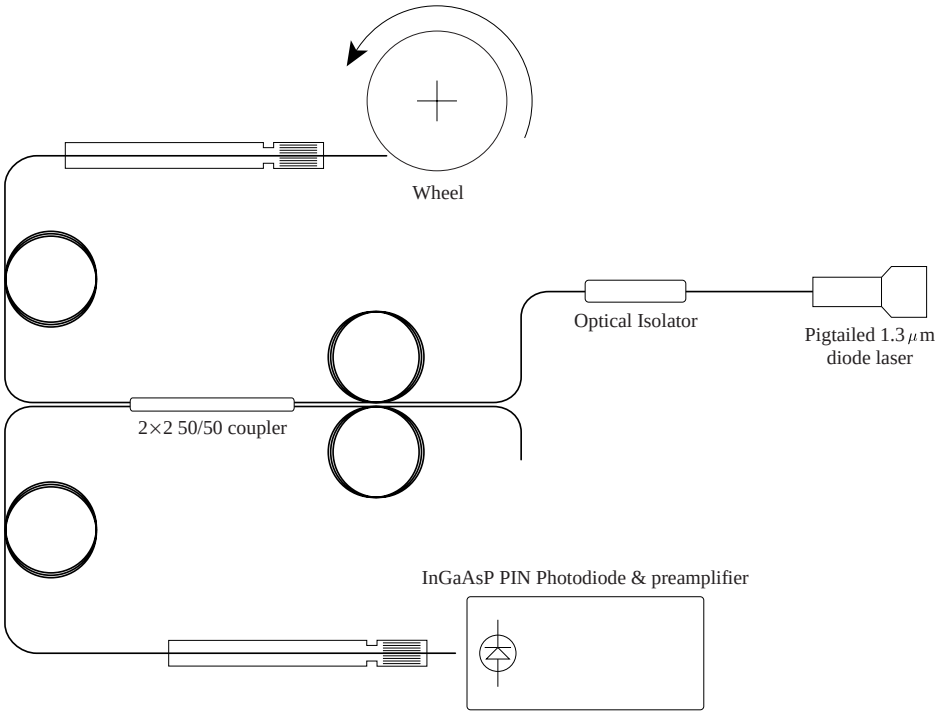

FIG. 1. Experimental setup for fiber-optic laser Doppler velocimeter.

The transmitter consists of a Hitachi ML776H11F InGaAsP distributed feedback, multi-quantum-well diode laser, of wavelength $\lambda = 1.31$ μm, nominal power $P_0 = 5$ mW, pigtailed to SMF28 single mode fiber. A low noise, constant current supply provided an injection current of 13.7 mA to the diode laser. No temperature control or external wavelength stabilization was employed.

A 2×2 fiber-optic coupler (CANSTAR DBS-02x02-131/155-50) with a 47.7%/52.3% coupling ratio was used to direct 52.3% of the laser output, or $P = 0.63$ mW, onto the rotating wheel target with reflectivity $R \sim 50\%$. In order to simulate a well-defined vibration frequency, the voltage drive to the wheel motor was sinusoidally modulated at a frequency near 20 Hz, resulting in a modulation of the wheel's angular velocity. The transceiver end of the fiber was cleaved normal to the fiber, positioned 1.38 cm below, and 1.37 cm laterally from the center of the 3.56 cm diameter wheel. The end of the fiber was thus $z = 2.4$ mm from the center of the illuminated spot on the wheel. For the data analyzed here, the

mean surface velocity component along the laser beam is about 10 cm/s, corresponding to a Doppler shift of approximately 200 kHz. The fiber nominally has a 9.3 μm mode field diameter and a numerical aperture of 0.13.

The backscattered signal is collected back into the fiber, transmitted through the other angle-cleaved, coupler input and is directed onto an Epitax ETX300T InGaAs PIN photodiode (quantum efficiency $\eta = 0.8$), which was AC-coupled to a low noise transimpedance amplifier (Analog Modules 711-4-4-AC). The detector/amplifier combination has a bandwidth $B = 1.5$ MHz, 2.9 pW/ $\mathrm{Hz}^{1/2}$ noise, and was operated with a transimpedance gain of approximately 3 V/μW. Voltage waveforms were collected at 2 megasample/s using a 12-bit, ± 1 V full scale A/D card (Adlink AD9812). Each sampled waveform consists of $216,136$ contiguous points, or 0.108 seconds of data (thus allowing about 2.4 vibrational periods).

The carrier to noise ratio (CNR) can be estimated as

$$\mathrm{CNR} = \frac{\mathrm{P}\,\lambda}{\mathrm{h\,c\,B}}\,\frac{\mathrm{A}}{\mathrm{z}^2}\,\mathrm{T}^2\mathrm{R}\,\eta\,\cos\theta\,, \tag{1}$$

where A is the fiber core cross-sectional area and θ is the scattering angle. With the parameters given above, and assuming 20% optical transmission efficiency T (including reflection losses, coupler losses and coupling to the detector) we can expect a CNR of about 20 dB. This can be considered the high CNR regime.

The purpose for selecting the rotating wheel geometry was both to provide a velocity offset for a reasonable Doppler carrier frequency (without frequency-shifting our local oscillator) and to provide a reasonable amount of speckle broadening for testing the robustness of the signal processing methods. We can determine the combined effect of speckle broadening and laser phase noise from the autocorrelation function of the experimental data, Figure 2. The envelope is roughly Gaussian with a $1/e$ width of 40 μs, for a coherence bandwidth of about 25 kHz.

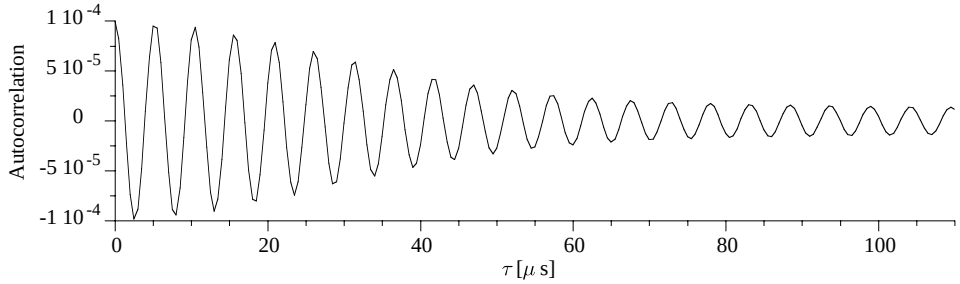

FIG. 2. Autocorrelation of the Doppler signal.

III. STATISTICAL INFERENCE IN SIGNAL PROCESSING

Since the development of the FFT, discrete Fourier transforms have become for many the first and last word in analyzing the frequency content of a signal. In many circumstances this use of discrete transforms is quite well justified. There are, however, many important instances where the FFT-based approach to spectral estimation is not optimal. Here we are primarily concerned with one such example, namely that of short time series. It is well known that the resolution of the discrete Fourier transforms is determined by the sample duration. When this interval only covers a few periods of the frequency of interest, it can be difficult to obtain good frequency estimates. This difficulty is further exacerbated by the leakage of the spectral response of the data windowing function[1] into the spectral range of the signal.

In the last fifteen years, a highly successful approach to data analysis based on statistical inference has emerged. These methods (which have been broadly labeled as Maximum Entropy) depart from traditional data-*processing* approaches in favor of the modeling of experiments. This distinction is more than merely pedantic and allows not only for a sound theoretical basis for estimation of various parameters of interest but also for assignment of confidence levels to these parameter estimates as well as for making relative quantitative assessments between competing models as to which is most consistent with the data.

Underlying all experiments is some model (hopefully encompassing a small number of parameters) that the experimenter believes will describe the experiment. The traditional approach to data analysis involves working "backwards" from the measured data to determine the parameters of interest in the model. Data analysis based on Bayesian statistical inference works "forwards" from the model and attempts to determine the model most statistically consistent with the data. Thus we ask the question "How likely is it that the observed data is a consequence of the model?" While on the surface, this approach appears to be nothing more than the fitting of parameters in the model to match the data, we shall see below that the statistical framework allows for much more. In particular, often there are parameters in our models that are essential for describing the data but otherwise do not contain physically relevant information. In a standard fitting approach, these uninteresting parameters would still nonetheless have to be included in along with "interesting" parameters. Moreover, it is not uncommon for these uninteresting parameters to out-number the interesting parameters, making the fitting procedure a great deal more work than if one could somehow consider only the physically relevant parameters. In the statistical approach it is possible

[1]Even if no window function is explicitly applied (a questionable practice in any event), there is an *implicit* windowing of the data that can not be avoided.

to marginalize (integrate out) these uninteresting parameters (often called nuisance parameters) leaving only the physically important parameters behind. Effectively this allows us to determine the values of the relevant parameters most consistent with the data knowing that the nuisance parameters will take on whatever values necessary to be consistent with the data. A germane example of nuisance parameters is the phase and quadrature amplitudes of a sinusoidal signal; very often we are only interested in the frequency. Bayesian methods provide the framework to consider the important parameters of our models while (effectively) ignoring the unimportant.

The goal of Bayesian[3] data analysis is to evaluate the conditional probability of values for the parameters in the model given the data and any prior information. Formally, through the use of Bayes' theorem we can write

$$P(\vartheta|\mathfrak{d},\mathfrak{I}) = \frac{P(\vartheta|\mathfrak{I})\,P(\mathfrak{d}|\vartheta,\mathfrak{I})}{P(\mathfrak{d}|\mathfrak{I})} \tag{2}$$

where ϑ denotes the set of parameters in the model, $\mathfrak{I}$ presents the prior information, and $\mathfrak{d}$ represents the measured data. Here $P(\vartheta|\mathfrak{d},\mathfrak{I})$ is the posterior probability of the model parameters given the data and prior information. This is the quantity of interest — the "best" values of the parameters are those which maximize this probability. The remaining terms in (2) have the following interpretations: $P(\vartheta|\mathfrak{I})$ is the probability of the parameters given only the prior information; $P(\mathfrak{d}|\mathfrak{I})$ is the probability of the data given the prior information;[2] and $P(\mathfrak{d}|\vartheta,\mathfrak{I})$ is the probability of the data given the parameters and prior information (this is often referred to as the likelihood of the data).

The prior probability of the parameters, $P(\vartheta|\mathfrak{I})$, is meant to summarize our knowledge of the model parameters *before* the experiment is performed. As long as there is sufficient data, the choice of this prior will have little influence on the final probability. It is generally accepted that a conservative approach is to choose a prior that represents "complete ignorance" regarding the parameter values. We will be more specific about the choice of the priors below when we examine the specific model representing our time series.

The likelihood of the data is perhaps conceptually the most straightforward part of the calculation. We denote our discrete set of N data samples as $\mathfrak{d} = \{d_j\}_{j=0}^{N-1}$, which correspond to the sample times $\{t_j\}_{j=0}^{N-1}$. The measured data is assumed to be described by a model $g(t;\vartheta)$ plus noise $\varepsilon(t)$, *i.e.*,

$$d_j = g(t_j;\vartheta) + \varepsilon(t_j), \qquad j = 0, 1, 2 \ldots N-1. \tag{3}$$

[2]For a given model and data set, this term can be considered a normalization constant. It is only of interest in comparing the relative probability of different models.

Typically we have little or no information regarding the measurement noise, yet we must somehow assign a prior probability to the noise. Traditionally this choice has been motivated by the principle of maximum entropy, that is, we assign the noise the least informative probability density, namely:

$$P(\varepsilon|\sigma,\mathfrak{I}) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\varepsilon^2/2\sigma^2}. \tag{4}$$

Under the assumption that the noise is uncorrelated, the probability that the measurement has the set of noise values $\mathfrak{e} = \{\varepsilon_j\}_{j=0}^{N-1}$ is

$$P(\mathfrak{e}|\sigma,\mathfrak{I}) \propto \prod_{j=0}^{N-1} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\varepsilon_j^2/(2\sigma^2)} \propto \sigma^{-N} \exp\left\{-\frac{1}{2\sigma^2}\sum_{j=0}^{N-1}(d_j - g(t_j;\vartheta))^2\right\}. \tag{5}$$

Combining this result with (2) we have

$$P(\vartheta|\mathfrak{d},\mathfrak{I}) \propto P(\vartheta|\mathfrak{I}) \times \sigma^{-N} \exp\left\{-\frac{1}{2\sigma^2}\sum_{j=0}^{N-1}(d_j - g(t_j;\vartheta))^2\right\}, \tag{6}$$

where the parameter set has been expanded to include the noise variance σ.

To proceed further it is useful to have a concrete example. In this work, we employ Bayesian analysis to extract the vibration spectrum from the time series of Doppler shifts. In this case we have both a relatively short time series (in terms of the vibration period) and a significant stochastic "noise" component in the signal (in part due to speckle broadening). While the determination of a single sinusoidal signal has become the canonical example of Bayesian data analysis, these methods are extremely versatile and have wide applicability. We take our data model to consist of a single harmonic frequency:

$$g(t;\omega,B_1,B_2) = B_1\,\cos\omega\,t + B_2\,\sin\omega\,t. \tag{7}$$

With this model, the posterior probability of the parameters becomes

$$P(\omega,B_1,B_2,\sigma|\mathfrak{d},\mathfrak{I}) \propto P(\vartheta|\mathfrak{I}) \times \sigma^{-N} \exp\left\{-\frac{1}{2\sigma^2}\sum_{j=0}^{N-1}(d_j - B_1\,\cos\omega\,t_j - B_2\,\sin\omega\,t_j)^2\right\}. \tag{8}$$

We will treat the phase and quadrature amplitudes as well as the noise variance as nuisance parameters and consider the marginal probability $P(\omega|\mathfrak{d},\mathfrak{I})$. We will take uniform priors for B_i and adopt the so-called Jeffreys prior, $1/\sigma$, for the noise variance.[3] Thus we have

$$P(\omega|\mathfrak{d},\mathfrak{I}) \propto \int d\sigma \int dB_1\, dB_2\, \sigma^{-(N+1)} \exp\left\{-\frac{1}{2\sigma^2}\sum_{j=0}^{N-1}(d_j - B_1\,\cos\omega\,t_j - B_2\,\sin\omega\,t_j)^2\right\}. \tag{9}$$

In this application, our data is uniformly sampled in time. Let Δt be the sampling interval and define $\Omega = \omega \Delta t$. The argument of the exponential in (9) can then be written as

$$-\frac{1}{2\sigma^2}\sum_{j=0}^{N-1}\left\{d_j^2 + B_1^2\cos^2 j\,\Omega + B_2^2\sin^2 j\,\Omega + B_1 B_2 \sin 2j\,\Omega - 2B_1 d_j \cos j\,\Omega - 2B_2 d_j \sin j\,\Omega\right\}$$
$$= N\overline{d^2} + B^T M B - 2B^T b, \tag{10}$$

where

$$B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \tag{11}$$

$$M = \begin{pmatrix} \sum_{j=0}^{N-1}\cos^2 j\,\Omega & \sum_{j=0}^{N-1}\sin 2j\,\Omega \\ \sum_{j=0}^{N-1}\sin 2j\,\Omega & \sum_{j=0}^{N-1}\sin^2 j\,\Omega \end{pmatrix}, \tag{12}$$

$$b = \begin{pmatrix} \sum_{j=0}^{N-1} d_j \cos j\,\Omega \\ \sum_{j=0}^{N-1} d_j \sin j\,\Omega \end{pmatrix}. \tag{13}$$

and $\overline{d^2}$ is the average of d_j^2 over the data set. It turns out that the sums in (12) can be computed in closed form, yielding

$$M = \begin{pmatrix} \frac{1}{2}\left[N + \frac{\sin N\Omega\,\cos(N-1)\Omega}{\sin\Omega}\right] & \frac{1}{2}\frac{\sin N\Omega\,\sin(N-1)\Omega}{\sin\Omega} \\ \frac{1}{2}\frac{\sin N\Omega\,\sin(N-1)\Omega}{\sin\Omega} & \frac{1}{2}\left[N - \frac{\sin N\Omega\,\cos(N-1)\Omega}{\sin\Omega}\right] \end{pmatrix}. \tag{14}$$

Note that for $0 < \Omega < \pi$, M has positive definite eigenvalues.

We can now write (9) as

$$P(\omega|\mathfrak{d},\mathfrak{I}) \propto \int_0^\infty d\sigma\;\sigma^{-(N+1)}\int_{-\infty}^{\infty} dB_1\,dB_2\;e^{-(N\overline{d^2}+B^T M B - 2B^T b)/2\sigma^2}. \tag{15}$$

We evaluate the Gaussian integral by introducing the change of variables $\widehat{B} = B - M^{-1}b$. Now $dB_1\,dB_2 = d\widehat{B}_1\,d\widehat{B}_2$ and we have

$$\int_{-\infty}^{\infty} dB_1\,dB_2\;e^{-(N\overline{d^2}+B^T M B - 2B^T b)/2\sigma^2} = e^{-(N\overline{d^2} - b^T M^{-1} b)/2\sigma^2}\int_{-\infty}^{\infty} d\widehat{B}_1\,d\widehat{B}_2\;e^{-\widehat{B}^T M \widehat{B}/2\sigma^2}. \tag{16}$$

Since M is symmetric, there exists an orthogonal matrix S such that $S^T\,M\,S = \mathrm{diag}(\lambda_1, \lambda_2)$. Let $u = S^T\,\widehat{B}$. Since $\det S = 1$, (16) becomes

$$\begin{aligned}\int_{-\infty}^{\infty} dB_1\,dB_2\; e^{-(N\overline{d^2}+B^TMB-2B^Tb)/2\sigma^2} &= e^{-(N\overline{d^2}-b^TM^{-1}b)/2\sigma^2}\int_{-\infty}^{\infty} du_1\,du_2\; e^{-u_1^2\lambda_1/2\sigma^2}e^{-u_2^2\lambda_2/2\sigma^2}\\ &= e^{-(N\overline{d^2}-b^TM^{-1}b)/2\sigma^2}\,\frac{2\pi\sigma^2}{\sqrt{\det M}}. \end{aligned} \tag{17}$$

This leaves only the integration over σ in (15), an expression which can be evaluated using a standard result:[5]

$$\int_0^{\infty} dx\; x^{\alpha-1}e^{-Cx} = \frac{\Gamma(\alpha)}{C^{\alpha}}\,. \tag{18}$$

Doing so gives

$$P(\omega|\mathfrak{d},\mathfrak{I}) \propto \frac{Q^{1-N/2}}{\sqrt{\det M}}\,, \tag{19}$$

where we have defined $Q = N\overline{d^2} - b^TM^{-1}b$.

We now have the marginal posterior probability as a function of ω alone; for any given data set, the value of ω that maximizes (19), which we will denote by ω_*, is the frequency that is most consistent with the data and the model. The values of parameters that have been marginalized can be estimated from the corresponding expectation values evaluated at $\omega = \omega_*$. The expectation value of any function φ of the model parameters is given by

$$\langle\varphi\rangle = \frac{\displaystyle\int d\sigma\int dB_1\,dB_2\;\varphi(\omega;\sigma,B_1,B_2)\,P(\vartheta|\mathfrak{d},\mathfrak{I})}{\displaystyle\int d\sigma\int dB_1\,dB_2\;P(\vartheta|\mathfrak{d},\mathfrak{I})}\,. \tag{20}$$

For example, we obtain an estimate of the phase and quadrature amplitudes from

$$\begin{aligned}\langle B\rangle\; &= \frac{\displaystyle\int d\widehat{B}_1\,d\widehat{B}_2\;\left(\widehat{B}+M^{-1}b\right)\,e^{-\widehat{B}^TM\widehat{B}/2\sigma^2}}{\displaystyle\int d\widehat{B}_1\,d\widehat{B}_2\;e^{-\widehat{B}^TM\widehat{B}/2\sigma^2}}\\ &= M^{-1}b\,.\end{aligned} \tag{21}$$

By a similar calculation we find

$$\left\langle B^TB\right\rangle\; = b^TM^{-2}b + 2\sigma^2\,\frac{N}{\det M} \tag{22}$$

and

$$\left\langle\sigma^2\right\rangle\; = \frac{Q}{N-4}\,. \tag{23}$$

From (22) we can estimate the signal-to-noise ratio

$$\mathrm{SNR} = \sqrt{2\frac{\mathrm{N}}{\det \mathrm{M}} + \frac{\langle \mathrm{B}^{\mathrm{T}}\mathrm{B}\rangle}{\langle\sigma^2\rangle}}\,, \tag{24}$$

where we have approximated σ^2 by $\langle\sigma^2\rangle$. Furthermore, we can estimate the uncertainty in ω_*. From (8) and (16) we can see that

$$P(\omega, \sigma|\mathfrak{d}, \mathfrak{I}) \propto e^{-Q/2\sigma^2}. \tag{25}$$

Expanding Q around ω_* we have

$$P(\omega, \sigma|\mathfrak{d}, \mathfrak{I}) \propto e^{-Q''(\omega_*)(\omega-\omega_*)^2/4\sigma^2}. \tag{26}$$

Taking the width of this distribution as a measure of the uncertainty in ω_* gives

$$\Delta\omega_* = \sqrt{\frac{2\,\langle\sigma^2\rangle}{Q''(\omega_*)}}, \tag{27}$$

where we have again approximated σ^2 by $\langle\sigma^2\rangle$.

IV. RESULTS

As mentioned above, here we apply Bayesian analysis to the problem of determining the vibration frequency which is manifest as a modulation of the Doppler shifted laser light. For our experimental arrangement, the average Doppler frequency is significantly higher than the imposed vibration frequency and is well resolved by the sampling rate. For these reasons, it is sufficient to determine the instantaneous Doppler shift using traditional FFT methods. We take the measured time series and split it into a sequence of "windows." After removing any DC component and applying a Bartlett window function, the power spectrum is estimated using the FFT and the frequency corresponding to the maximum

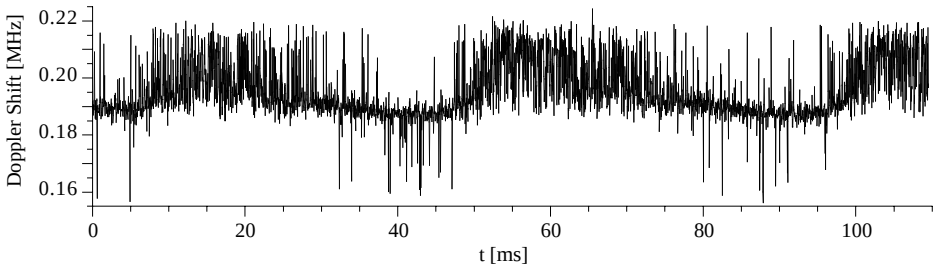

FIG. 3. Doppler shift determined using 32 μs window.

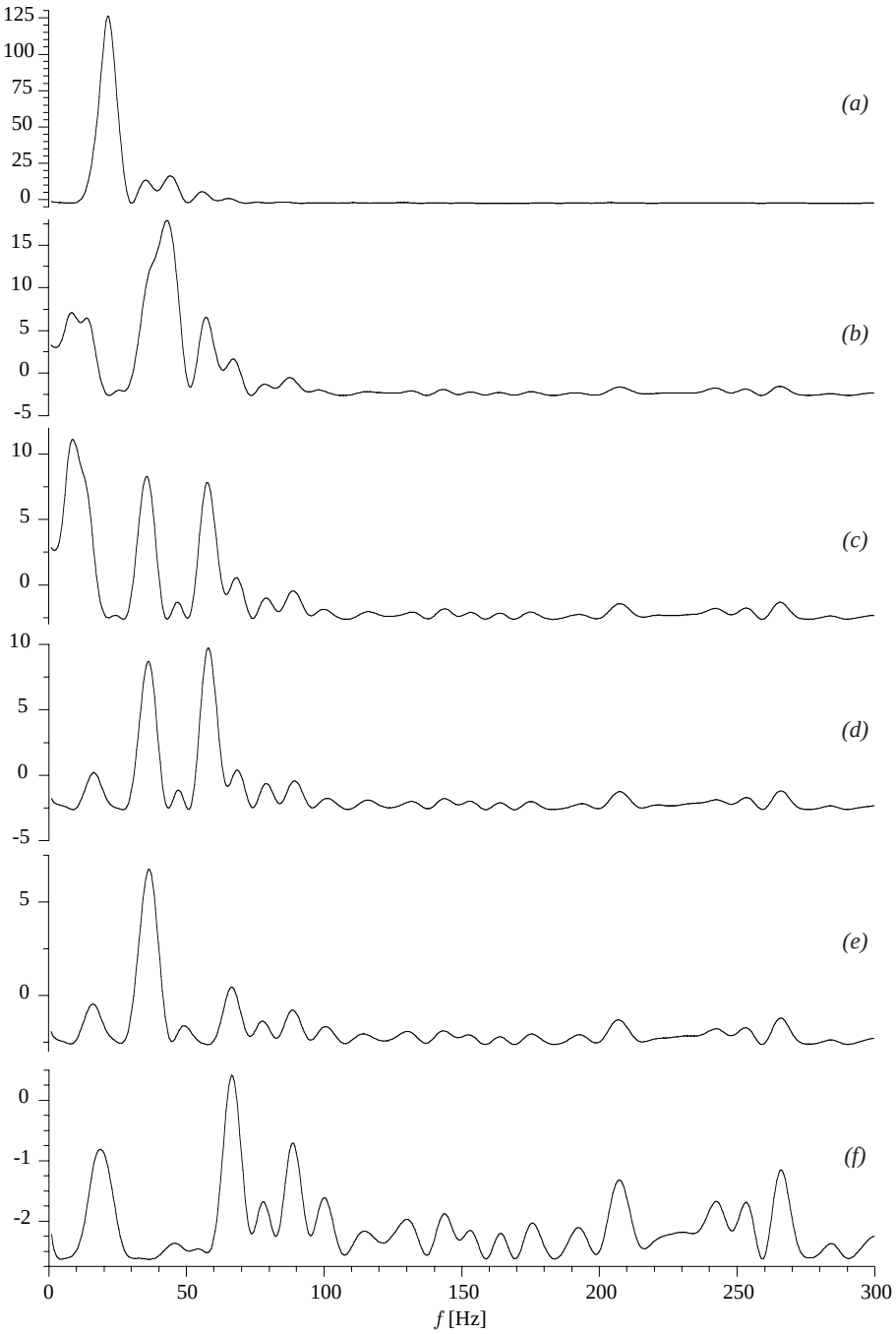

FIG. 4. Plot of $\log_{10} P(\omega|\mathfrak{e},\mathfrak{I})$ as given by (19) for the Doppler signal obtained from the 128 μs window *vs.* f ($\omega = 2\pi f$): (a) for the original signal; (b) for the residual after the removal of the dominant frequency in (a). Each of the remaining panels shows probability after removing the dominant frequency of the previous panel.

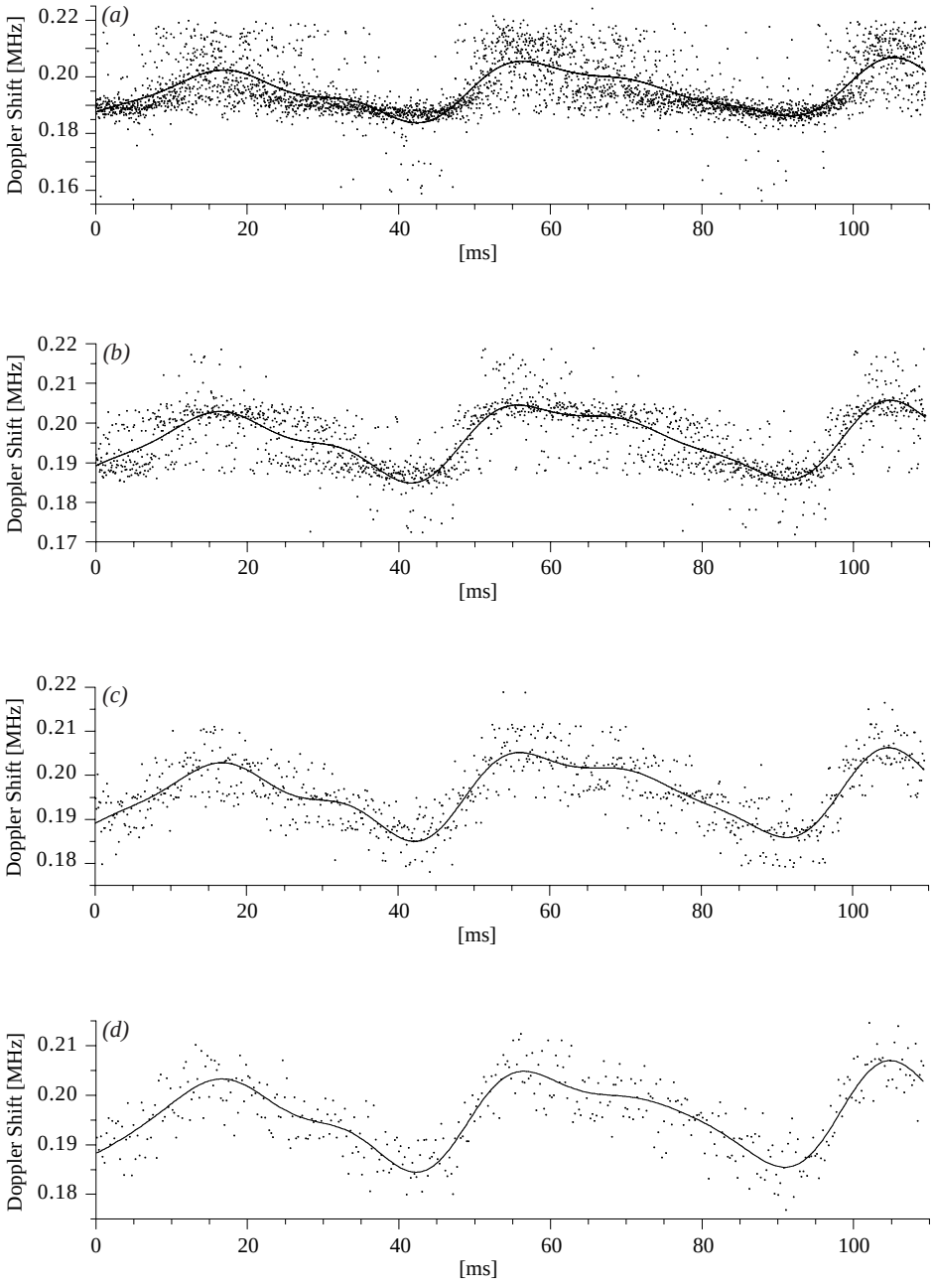

FIG. 5. Best fit (solid line) and raw time series (dots) using (*a*) 32 μs, (*b*) 64 μs, (*c*) 128 μs, and (*d*) 256 μs windows. The best fits correspond to the parameters shown in Table (I).

32 μs window

f [Hz]	B_1 [kHz]	B_2 [kHz]	$\lvert B\rvert$ [kHz]
8.67 ± 0.51	-1.857	-0.711	1.989
22.08 ± 0.11	-5.733	5.850	8.191
36.30 ± 0.57	0.925	-1.216	1.528
42.73 ± 0.49	-2.257	0.966	2.455
57.37 ± 0.41	1.913	0.352	1.945

128 μs window

f [Hz]	B_1 [kHz]	B_2 [kHz]	$\lvert B\rvert$ [kHz]
8.88 ± 0.68	-1.726	-0.882	1.938
21.70 ± 0.16	-5.124	5.829	7.761
36.64 ± 0.73	0.819	-1.246	1.491
43.14 ± 0.64	-2.244	1.077	2.489
58.03 ± 0.56	1.730	0.356	1.766

64 μs window

f [Hz]	B_1 [kHz]	B_2 [kHz]	$\lvert B\rvert$ [kHz]
8.79 ± 0.56	-1.718	-0.497	1.788
21.70 ± 0.13	-5.021	6.004	7.826
35.52 ± 0.74	0.261	-1.347	1.372
42.43 ± 0.67	-1.475	1.682	2.237
57.74 ± 0.50	1.743	-0.196	1.754

256 μs window

f [Hz]	B_1 [kHz]	B_2 [kHz]	$\lvert B\rvert$ [kHz]
9.14 ± 1.2	-0.911	-0.945	1.313
$21.87 \pm .21$	-5.562	5.543	7.853
$35.88 \pm .77$	0.639	-1.713	1.828
$43.42 \pm .70$	-2.657	0.906	2.808
58.31 ± 1.1	1.246	0.428	1.317

TABLE I. Frequency and amplitude estimates using various window sizes. The frequencies are extracted from the signal sequentially as described in the text and the uncertainties are computed using (27) and the amplitudes are obtained using (22).

power is determined by quadratic interpolation. This process yields a time series of Doppler frequencies with a sample interval corresponding to the window length. It is this derived time series that we analyze using the Bayesian methods described above.

We present results for four different length windows: 64, 128, 256, and 512 samples corresponding to 32 μs, 64 μs, 128 μs and 256 μs, respectively. Figure 3 shows the time series of Doppler shifts determined using the 32 μs window. The time series clearly contains a significant stochastic component, however by eye one can discern a periodic structure. In this application $\mathfrak{e}$ does not represent measurement "noise" but rather the stochastic component of the signal, *i.e.,* we are considering our signal to be composed of a deterministic part which we model by g and a stochastic part, $\mathfrak{e}$, which we assume has a Gaussian distribution. The statistical arguments are the same as in the case where $\mathfrak{e}$ corresponds to experimental noise, but the philosophy is slightly different.

In Figure 4(a), we plot the base 10 logarithm of the posterior probability, (19), as a function of frequency f for the 128 μs window. As is typical with Bayesian analysis, the

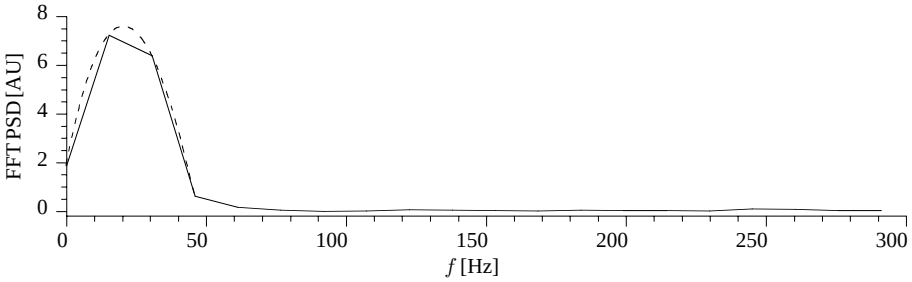

FIG. 6. Vibration power spectral density estimated using FFT (solid line) and cubic interpolation of the peak (dashed line). The PSD has a maximum at $f = 20.6$ Hz with a HWHM of 17.5 Hz.

peak in the probability is vastly above the background level. Such sharp peaks give rise to precise estimates of the parameters. We can also see that the probability plot suggests the presence of other frequencies in addition to that responsible for the main peak. Although our model only contains a single frequency, we can analyze multi-harmonic time series by iteratively removing the frequency corresponding to the peak in the probability and then re-analyzing the residual.[3] A more complete approach would be to carry out a full multi-mode analysis, using $P(\mathfrak{d}|\mathfrak{I})$ to select the optimal model. Such an analysis is significantly more computationally intensive than the recursive method used here. As we will see below, the results from this simpler approach are of sufficiently high quality that the more complex multi-mode method is not justified by our data.

Shown in Figure 4(b)–(f) are posterior probabilities of the residual after removal of subsequent frequencies. Notice how the vertical scale changes drastically; as each frequency is removed the peak in the probability corresponding to the "next" frequency is much smaller relative to the background. We attribute these additional frequencies to (i) harmonics of the drive modulation frequency due to the "accelerate and coast" effect of the drive modulation and (ii) motor "cogging" effects due to the low chopper wheel velocity. When we reach panel (f), the probability no longer contains any dominant features. The results of this procedure for all four window sizes are summarized in Table I. For the three largest windows the frequency estimates are in complete agreement within their respective one-σ uncertainties. The results from the 32 μs window are also in agreement with the other cases except for the frequency of the mode near 22 Hz. We attribute this (small) discrepancy to the fact that the window length of 32 μs is *smaller* than the speckle induced decoherence time of approximately 40 μs. Thus for this short window, the stochastic component may deviate significantly for the Gaussian distribution that it achieves over longer time intervals.

[3]The approach is known to work well with sinusoidal signals but can be disastrous in other circumstances.[6]

This has the effect of making the model not quite correct, *i.e.,* our assumptions about the "noise" statistics are not valid. It is then not surprising that we get some disagreement with the cases where the model more closely matches the data.

In Figure 6, we show the FFT estimate of the power spectral density of the time series produced using the 256 μs window. The PSD has a maximum at $f = 20.6$ Hz with a half-width-half-max of 17.5 Hz. This maximum value is approximately 1.2 Hz below the average frequency for this mode determined from the Bayesian analysis. Clearly the PSD is incapable of resolving the multiple frequencies contained in the vibration signal. While the peak (after interpolation) yields approximately the same value for the best vibration frequency, the significant width of the spectrum (due to the small number of periods contained in the time series) makes an accurate estimate of the frequency difficult. Contrary to popular belief, "enlarging" the data set by zero-padding does not improve the resolution of the FFT but merely interpolates in power spectrum between the frequencies of the shorter data set. The width of this power spectrum estimate is a fundamental limitation of using the FFT and cannot be side-stepped.

V. CONCLUSIONS AND FUTURE WORK

Here we have demonstrated the power of Bayesian analysis when dealing with short time series containing a significant stochastic component. We have shown that the Bayesian approach yields much more precise frequency estimates than is possible when using the FFT as a spectral estimator. We have also seen that it is important that the model accurately fit the data as this is part of the "prior information." As with the priors for the noise or the other model parameters, choosing a poor model can color the results of the analysis. Ultimately we are asking for the parameter values which, given the model, are most consistent with the data. The Bayesian framework provides a means of selecting between competing models and any complete analysis should include consideration of multiple models. Here we considered only a single sinusoid model since we were most concerned with demonstrating the use of Bayesian methods as a "proof-of-principle." This deficiency notwithstanding, the Bayesian approach yields a frequency estimate that is accurate to approximately 1%. Given the short time series, this is nonetheless quite impressive as it represents a significant improvement (by nearly two orders of magnitude) upon the estimate obtained from the FFT.

This method will have the greatest impact in situations where measurement dwell-time is at a premium, such as vibrational imaging [7] where many points across the surface of an object must be monitored sequentially during a measurement period, or in situations where the vibrational character is rapidly changing and dwell-time must be limited. If the

experimental circumstances were such that there was not such a large separation between the Doppler frequency and the vibration frequency, then using such large windows for determining the Doppler frequency would not be appropriate. In such a case we could also employ Bayesian analysis to determine the Doppler shift. While not necessary with the data considered here, using Bayesian methods should allow for resolving vibration frequencies that are a significant fraction of the Doppler frequency.

In future investigations we plan to evaluate the performance of statistical signal processing methods applied to laser vibrometry in the low to moderate CNR regime (≈ 0 dB) and with more realistic vibrational signals. The focus of this preliminary investigation was on determining the vibrational frequency with a limited amount of data corrupted by significant phase and frequency noise. Future investigations will also focus on the ability of the method to detect small vibrational amplitudes in comparison with other processing methods.

ACKNOWLEDGMENTS

BAS gratefully acknowledges many useful discussions with A. Charman on statistical signal processing.

REFERENCES

* Correpsonding author email: Walter.F.Buell@aero.org.

[1] A. L. Kachelmyer and K. I. Schultz, Linc. Lab. J. **8**, 3 (1995).

[2] T. D. Cole and A. S. El-Dinary, in *Applied Laser Radar Technology, Proc. SPIE*, edited by G. W. Kamerman and W. E. Keicher (SPIE, Bellingham, WA, 1993), Vol. 1936, pp. 90–103.

[3] G. L. Bretthorst, *Bayesian Spectrum Analysis and Parameter Estimation*, *Lecture Notes in Statistics, Volume 48* (Springer-Verlag, New York, 1988).

[4] R. W. Farley and Y. C. Chan (unpublished).

[5] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Vol. 55 of *Applied Mathematics Series* (National Bureau of Standards, Washington, 1964), reprinted by Dover Publications, New York, 1968.

[6] J. J. K. Ó Ruanaidh and W. J. Fitzgerald, *Numerical Bayesian Methods Applied to Signal Processing* (Springer, New York, 1996).

[7] P. Lutzmann, R. Frank, and R. Ebert, in *Laser Radar Technology and Applications V, Proc. SPIE*, edited by G. W. Kamerman, U. N. Singh, C. Werner, and V. V. Molebny (SPIE, Bellingham, WA, 2000), Vol. 4035, pp. 436–443.