Insights on the Bayesian spectral density method for operational modal analysis

Siu-Kui Au[1]
Center for Engineering Dynamics and Institute for Risk and Uncertainty
University of Liverpool

ABSTRACT

This paper presents a study on the Bayesian spectral density method for operational modal analysis, i.e., modal identification using output-only ambient vibration data. The method makes Bayesian inference of the modal properties by using the sample power spectral density (PSD) matrix averaged over independent sets of ambient data. In the typical case with a single set of consecutive time history of data, it is divided into non-overlapping segments and they are assumed to be independent. This study is motivated by a recent paper that reveals a mathematical equivalence of the method with the Bayesian FFT method. The latter is a more fundamental approach based directly on the FFT of the ambient data without averaging concepts or the independent segment assumption. This study shows that the equivalence does not hold in reality when the segment length is not long (compared to the natural period), because the theoretical long data asymptotic distribution of the PSD matrix can deviate significantly from reality. A single time history can be considered sufficiently long for the Bayesian FFT method but the same need not be true for the Bayesian PSD method, depending on the number of segments. Violating long data asymptotics leads to bias in the modal identification results, primarily as a result of spectral leakage.

Keywords: Bayesian FFT method, Bayesian spectral density method, ambient modal identification, operational modal analysis, spectral leakage

[1] Corresponding author. Institute for Risk and Uncertainty and Centre for Engineering Dynamics, University of Liverpool. Harrison Hughes Building, Brownlow Hill, L69 3GH. Tel: +44 (0)151 794 5217. E-mail: siukuiau@liverpool.ac.uk

1 Introduction

Ambient modal identification, often known as ‘operational modal analysis’, aims at identifying the modal properties (e.g., natural frequency, damping ratios, mode shapes) of a constructed structure using (output-only) measured vibration data [1][2]. The loading to the structure is not known but is assumed to be broadband random so that statistical characteristics of the response can be assumed to be attributed to the modal properties rather than the loading. Stochastic subspace identification [3] and frequency domain decomposition [4] are two conventional methods with variants in the time and frequency domain, respectively. In the absence of loading information, the uncertainty associated with the identified parameters can be significant. This calls for a rational approach for quantifying the identification uncertainty of modal parameters. In a classical statistical approach, also known as ‘frequentist’ approach, identification uncertainty is measured as the variability of the estimate over repeated experiments. Methods for calculating the uncertainty of modal parameters in this context have been developed [5].

A Bayesian system identification approach [6][7][8] provides a fundamental means for identifying the modal parameters consistent with probability logic and modeling assumptions. Bayesian approaches for ambient modal identification have been formulated in a series of papers by Yuen & Katafygiotis, chronologically, in the time domain [9], in the frequency domain based on the power spectral density (PSD) matrix [10], and in the frequency domain based on the Fast Fourier Transform (FFT) of the ambient data [11]. Judging solely on modeling assumptions, the time domain approach is less preferred to the frequency domain approaches because it inevitably involves imposing broadband assumption on the excitation and the prediction error (e.g., channel noise) over the whole sampling spectrum, i.e., up to the Nyquist frequency, which is unlikely to be valid unless the data has been filtered [12]. The PSD method is also less preferred to the FFT method because it involves averaging concepts and requires independent data sets.

In a Bayesian context, the identification result of the modal parameters is encapsulated in their joint posterior (i.e., given data) probability density function (PDF). With sufficient data the modal parameters are ‘globally identifiable’ [13], in the sense that their posterior PDF has a single regular peak. In this case the posterior PDF can be approximated by a Gaussian PDF

centered at the posterior most probable value (MPV). Mathematically, the MPV minimizes the negative log-likelihood function (NLLF). The covariance matrix of the Gaussian PDF, which reflects the identification uncertainty of the modal parameters, is equal to the inverse of the Hessian of the NLLF at the MPV. Determining the MPV by brute-force numerical optimization of the NLLF is computationally prohibitive, primarily because the number of parameters grows with the number of DOFs and is therefore typically large. This undermined the application of the aforementioned Bayesian approaches in their original form. By a mathematical reformulation of the NLLF of the FFT method, fast algorithms for computing the posterior MPV and covariance matrix were later developed, for well-separated modes [14], multiple (possibly close) modes [15][16] and well-separated modes from multiple setups [17][18]. These developments allowed practical application of the Bayesian FFT method, e.g., [19][20][21].

Following similar techniques used in developing the fast algorithms for the FFT method, an algorithm for determining the posterior statistics (MPV and covariance matrix) in the PSD method has been recently presented for general multiple modes [22]. The method is approximate because the MPV of the 'spectral parameters' (i.e., those other than the mode shapes) minimizes a surrogate NLLF formulated based on the trace of the PSD matrix as data rather than the PSD matrix itself. Nevertheless the paper suggested an interesting similarity of the mathematical structure of the PSD and FFT method. Motivated along this direction, this paper presents some insights on the PSD method and how it is related to the FFT method. For a single set of ambient data, which is typically the case in practice, it is shown that the PSD method is mathematically equivalent to dividing the data into non-overlapping segments, assuming that they are independent and making inference collectively based on their FFTs. In reality, however, the PSD method is more vulnerable to modeling error when the number of segments is not small.

This paper is organized as follows. We first present the Bayesian FFT method which provides the fundamental approach for making inference about the modal parameters making full use of the information within the selected frequency band. The Bayesian PSD method is then presented and its underlying assumptions are clarified. The mathematical equivalence of

these two approaches is then established, followed by a discussion. The relevant issues are illustrated with numerical examples based on synthetic data.

2 Bayesian FFT method

We first give a brief introduction of the Bayesian FFT method; a recent review can be found in [12]. Let $\{\hat{\ddot{\mathbf{x}}}_j \in R^n\}_{j=1}^N$ denote the measured digital time history of acceleration response at n DOFs of the subject structure. The method makes use of the FFT of the measured response on a selected resonance frequency band as data to formulate the posterior PDF of modal parameters. The FFT is defined as

$$\hat{\mathcal{F}}_k = \sqrt{\frac{2\Delta t}{N}} \sum_{j=1}^{N} \hat{\ddot{\mathbf{x}}}_j e^{-2\pi \mathbf{i}(j-1)(k-1)/N} \qquad k = 1,\ldots,N \qquad (1)$$

which corresponds to frequency $\mathrm{f}_k = (k-1)/N\Delta t$ up to the Nyquist frequency ($1/2\Delta t$ Hz); Δt is the sampling interval in seconds. The FFT of the measured data is modeled as

$$\hat{\mathcal{F}}_k = \sum_{i=1}^{m} \boldsymbol{\varphi}_i h_{ik} p_{ik} + \boldsymbol{\varepsilon}_k \qquad (2)$$

where $\{\boldsymbol{\varphi}_i \in R^n\}_{i=1}^m$ are the mode shape vectors of the m contributing modes in the selected frequency band; $\boldsymbol{\varepsilon}_k \in C^n$ is the FFT of the prediction error (e.g., channel noise), assumed to have a constant PSD matrix $S_e \mathbf{I}_n$ within the selected frequency band; $\mathbf{I}_n$ denotes the identity matrix of dimension n; $\{p_{ik}\}_{i=1}^m$ are the FFT at frequency $\mathrm{f_k}$ of the modal forces, assumed to have a constant PSD matrix $\mathbf{S} \in C^{n \times n}$ within the selected frequency band; and

$$h_{ik} = [(\beta_{ik}^2 - 1) + (2\zeta_i \beta_{ik})\mathbf{i}]^{-1} \qquad (3)$$

is the (complex-valued) transfer function between the i-th modal force and modal acceleration; $\beta_{ik} = \mathrm{f}_k / f_i$ is a frequency ratio of the FFT frequency f_k to the natural frequency f_i.

In the above context, the set of modal parameters to be identified is given by

$$\boldsymbol{\theta} = \{f_i, \zeta_i, \boldsymbol{\varphi}_i : i = 1,\ldots,m; \mathbf{S}, S_e\} \qquad (4)$$

with the constraints that $\mathbf{S}$ is a Hermitian matrix and the mode shapes are normalized to have unit norm, i.e., $\boldsymbol{\varphi}_i^T \boldsymbol{\varphi}_i = 1$ ($i = 1,...,m$). Assuming a uniform (constant) prior distribution for $\boldsymbol{\theta}$, its posterior PDF can be written in a compact form by [15]

$$p(\boldsymbol{\theta} \mid \{\hat{\mathcal{F}}_k\}) \propto \exp[-L(\boldsymbol{\theta})] \tag{5}$$

where

$$L(\boldsymbol{\theta}) = nN_f \ln \pi + \sum_k \ln \det \mathbf{E}_k + \sum_k \hat{\mathcal{F}}_k^* \mathbf{E}_k^{-1} \hat{\mathcal{F}}_k \tag{6}$$

and

$$\mathbf{E}_k(\boldsymbol{\theta}) = E[\mathcal{F}_k \mathcal{F}_k^* \mid \boldsymbol{\theta}] = \boldsymbol{\Phi} \mathbf{H}_k \boldsymbol{\Phi}^T + S_e \mathbf{I}_n \tag{7}$$

is the theoretical PSD matrix of the data. The asterisk '*' denotes a complex conjugate transpose.

The NLLF in (6) and hence the posterior PDF assumes that the data duration is long compared to the natural periods of the modes in the selected frequency band so that the FFTs $\{\mathcal{F}_k\}$ at different frequencies are asymptotically independent and they are jointly complex Gaussian, each with a covariance matrix given by $\mathbf{E}_k(\boldsymbol{\theta}) = E[\mathcal{F}_k \mathcal{F}_k^* \mid \boldsymbol{\theta}]$. In the limit this matrix becomes the theoretical PSD matrix in (7). This is an important aspect that will be revisited later when the PSD method is discussed.

Modal identification problems with sufficient data are globally identifiable, in the sense that the posterior PDF of $\boldsymbol{\theta}$ has a single peak in the parameter space. In this case, using a second order Taylor expansion of the NLLF, the posterior PDF can be approximated by a Gaussian PDF centered at the most probable value (MPV) and with a covariance matrix. Mathematically, the MPV minimizes the NLLF and the covariance matrix is equal to the inverse of the Hessian of the NLLF evaluated at the MPV.

The Bayesian FFT method is fundamental in the sense that inference is made directly based on the FFT without averaging concepts. The FFT is not a statistical proxy as it has a one-one correspondence with the time domain data. There is no loss in information by making inference based on the FFT rather than the time domain data directly. Using the FFT data on

a selected frequency band, usually around the resonance peak of the modes of interest, allows one to play a balance between the information used for identification (the wider the band the better) and modeling error risk (the wider the band the higher the risk).

3 Bayesian PSD method

The Bayesian PSD method was developed earlier than the Bayesian FFT method. It makes Bayesian inference based on the sample PSD matrix of the ambient data averaged over M (say) independent sets of data. In real situations where only one consecutive time series of ambient data is available, the M sets of data are obtained by dividing the original data set of duration T_d into M non-overlapping segments (each of duration T_d / M) and they are assumed to be independent. Let $\{\hat{\ddot{\mathbf{x}}}_j^{(r)}\}_{j=1}^{N/M}$ denote the r -th segment ($r = 1,..., M$):

$$\hat{\ddot{\mathbf{x}}}_j^{(r)} = \hat{\ddot{\mathbf{x}}}_{j+(r-1)N/M} \qquad j = 1,..., N/M \tag{8}$$

Let $\{\hat{F}_{k'}^{(r)}\}$ denote the FFT of $\{\hat{\ddot{\mathbf{x}}}_j^{(r)}\}_{j=1}^{N/M}$ in the selected frequency band, i.e.,

$$\hat{F}_{k'}^{(r)} = \sqrt{\frac{2M\Delta t}{N}} \sum_{j=1}^{N/M} \hat{\ddot{x}}_j^{(r)} e^{-2\pi \mathbf{i} M(j-1)(k'-1)/N} \qquad k' = 1,..., N/M \tag{9}$$

which now corresponds to the frequency $\mathrm{f}_{k'} = M(k'-1)/N\Delta t$ (Hz). A different symbol k' (instead of k) is used for the frequency index to reflect the fact that the FFT now has a lower frequency resolution than the FFT in (1), as the frequency interval is now $M/N\Delta t$ (Hz) instead of $1/N\Delta t$. The sample PSD matrix averaged over M segments is given by

$$\hat{\mathbf{E}}_{k'} = \frac{1}{M} \sum_{r=1}^{M} \hat{F}_{k'}^{(r)} \hat{F}_{k'}^{(r)*} \tag{10}$$

The PSD method takes the collection $\{\hat{\mathbf{E}}_{k'}\}$ in the selected frequency band as the 'data' for making inference about the modal parameters. In formulating the likelihood function, it makes use of the theoretical result that, for $M \geq n$ and assuming large segment length T_d / M , the PDF of $\hat{\mathbf{E}}_{k'}$ for given $\boldsymbol{\theta}$ has a 'complex Wishart distribution' of dimension n and with M degrees of freedom [10][23][24]. This PDF is given by

$$p(\hat{\mathbf{E}}_{k'} \mid \boldsymbol{\theta}) = \frac{(\det \hat{\mathbf{E}}_{k'})^{M-n} \pi^{-n(n-1)/2} M^{nM}}{[\det \mathbf{E}_{k'}(\boldsymbol{\theta})]^{M} \prod_{i=1}^{n} (M-i)!} \exp\left\{-M \ tr[\mathbf{E}_{k'}(\boldsymbol{\theta})^{-1} \hat{\mathbf{E}}_{k'}]\right\} \qquad (M \geq n) \tag{11}$$

where $\mathbf{E}_{k'}(\boldsymbol{\theta}) = E[\hat{\mathbf{E}}_{k'} \mid \boldsymbol{\theta}]$ is given by (7) with k' substituting k; $tr[\cdot]$ denotes the trace of the argument matrix, i.e., the sum of diagonal entries. The likelihood function in the PSD method is the product of $p(\hat{\mathbf{E}}_{k'} \mid \boldsymbol{\theta})$ for different frequencies in the selected band, i.e.,

$$p(\{\hat{\mathbf{E}}_{k'}\} \mid \boldsymbol{\theta}) = \prod_{k'} \frac{(\det \hat{\mathbf{E}}_{k'})^{M-n} \pi^{-n(n-1)/2} M^{nM}}{[\det \mathbf{E}_{k'}(\boldsymbol{\theta})]^{M} \prod_{i=1}^{n} (M-i)!} \exp\left\{-M \ tr[\mathbf{E}_{k'}(\boldsymbol{\theta})^{-1} \hat{\mathbf{E}}_{k'}]\right\} \qquad (M \geq n) \tag{12}$$

Assuming a uniform (constant) prior distribution for $\boldsymbol{\theta}$, its posterior PDF taking $\{\hat{\mathbf{E}}_{k'}\}$ as data is given by

$$p(\boldsymbol{\theta} \mid \{\hat{\mathbf{E}}_{k'}\}) \propto p(\{\hat{\mathbf{E}}_{k'}\} \mid \boldsymbol{\theta}) \propto e^{-L_{PSD}(\boldsymbol{\theta})} \tag{13}$$

where

$$L_{PSD}(\boldsymbol{\theta}) = M \sum_{k'} \ln \det \mathbf{E}_{k'}(\boldsymbol{\theta}) + M \sum_{k'} \mathrm{tr}[\mathbf{E}_{k'}(\boldsymbol{\theta})^{-1} \hat{\mathbf{E}}_{k'}] \qquad (M \geq n) \tag{14}$$

is the NLLF for the PSD method.

As mentioned before, the NLLF in (14) and hence the posterior PDF assumes that the segment length T_d / M is large compared to the natural periods of the modes in the selected frequency band, so that each $\hat{\mathbf{E}}_{k'}$ has a complex Wishart distribution and the $\hat{\mathbf{E}}_{k'}$s at different frequencies are asymptotically independent. The additional condition $M \geq n$ ensures that $\hat{\mathbf{E}}_{k'} \in C^{n \times n}$ as a random matrix is almost surely non-singular because its rank is $\min(n, M)$.

4 Apparent mathematical equivalence

In this section we show that when dealing with a single set of consecutive time-domain data, which is typically the case in practice, the PSD method (for $M \geq n$) is mathematically equivalent to making inference collectively based on the FFTs of the M non-overlapping

segments by assuming that they are independent and have large segment length T_d / M. Specifically, we shall show that the NLLF derived under this context is identical to that of the PSD method. In later sections we shall argue that in reality this equivalence does not hold (or not useful in practice) because T_d / M need not be large even if T_d is large.

The apparent equivalence is based on the observation that the trace term in the NLLF (14) of the PSD method can be written as a quadratic form similar to that in the NLLF (6) of the FFT method. This in turn is based on the 'cyclic' property of the trace function, which was also used in developing the algorithm in [22]. Specifically, for any complex matrices X, Y, Z of appropriate dimension,

$$tr[XYZ] = tr[ZXY] \tag{15}$$

Substituting $\hat{\mathbf{E}}_{k'}$ from (10) and using this identity, the second term in (14) becomes

$$\begin{aligned}
&M \sum_{k'} \mathrm{tr}[\mathbf{E}_{k'}(\boldsymbol{\theta})^{-1} \hat{\mathbf{E}}_{k'}] \\
&= M \sum_{k'} tr[\mathbf{E}_{k'}(\boldsymbol{\theta})^{-1} \frac{1}{M} \sum_{r=1}^{M} \hat{\mathcal{F}}_{k'}^{(r)} \hat{\mathcal{F}}_{k'}^{(r)*}] \\
&= \sum_{r=1}^{M} \sum_{k'} tr[\mathbf{E}_{k'}(\boldsymbol{\theta})^{-1} \hat{\mathcal{F}}_{k'}^{(r)} \hat{\mathcal{F}}_{k'}^{(r)*}] \\
&= \sum_{r=1}^{M} \sum_{k'} tr[\hat{\mathcal{F}}_{k'}^{(r)*} \mathbf{E}_{k'}(\boldsymbol{\theta})^{-1} \hat{\mathcal{F}}_{k'}^{(r)}] \\
&= \sum_{r=1}^{M} \sum_{k'} \hat{\mathcal{F}}_{k'}^{(r)*} \mathbf{E}_{k'}(\boldsymbol{\theta})^{-1} \hat{\mathcal{F}}_{k'}^{(r)}
\end{aligned} \tag{16}$$

where the second equality follows because the trace function is linear (trace of sum is equal to sum of trace); the last equality follows because $\hat{\mathcal{F}}_{k'}^{(r)*} \mathbf{E}_{k'}(\boldsymbol{\theta})^{-1} \hat{\mathcal{F}}_{k'}^{(r)}$ is a scalar. Substituting (16) into (14),

$$L_{PSD}(\boldsymbol{\theta}) = M \sum_{k'} \ln \det \mathbf{E}_{k'}(\boldsymbol{\theta}) + \sum_{r=1}^{M} \sum_{k'} \hat{\mathcal{F}}_{k'}^{(r)*} \mathbf{E}_{k'}(\boldsymbol{\theta})^{-1} \hat{\mathcal{F}}_{k'}^{(r)} \qquad (M \geq n) \tag{17}$$

On the other hand, consider making Bayesian inference using the FFTs calculated from the M segments of data. Assume that they are independent and the segment length T_d / M is

sufficiently long. The likelihood function in the present context is equal to the product of those of the individual segments because they are assumed independent:

$$p(\{\hat{\mathcal{F}}_{k'}^{(1)}\},\{\hat{\mathcal{F}}_{k'}^{(2)}\},...,\{\hat{\mathcal{F}}_{k'}^{(M)}\}\,|\,\boldsymbol{\theta}) = \prod_{r=1}^{M} p(\{\hat{\mathcal{F}}_{k'}^{(r)}\}\,|\,\boldsymbol{\theta}) \tag{18}$$

The likelihood function of each segment is given by

$$p(\{\hat{\mathcal{F}}_{k'}^{(r)}\}\,|\,\boldsymbol{\theta}) \propto e^{-L_r(\boldsymbol{\theta})} \tag{19}$$

with the NLLF being, based on (6) and omitting the constant,

$$L_r(\boldsymbol{\theta}) = \sum_{k'} \ln\det \mathbf{E}_{k'}(\boldsymbol{\theta}) + \sum_{k'} \hat{\mathcal{F}}_{k'}^{(r)*} \mathbf{E}_{k'}(\boldsymbol{\theta})^{-1} \hat{\mathcal{F}}_{k'}^{(r)} \tag{20}$$

The likelihood function in (18) is then given by

$$p(\{\hat{\mathcal{F}}_{k'}^{(1)}\},\{\hat{\mathcal{F}}_{k'}^{(2)}\},...,\{\hat{\mathcal{F}}_{k'}^{(M)}\}\,|\,\boldsymbol{\theta}) \propto \exp\left[-\sum_{r=1}^{M} L_r(\boldsymbol{\theta})\right] \tag{21}$$

where the NLLF is

$$\sum_{r=1}^{M} L_r(\boldsymbol{\theta}) = M \sum_{k'} \ln\det \mathbf{E}_{k'}(\boldsymbol{\theta}) + \sum_{r=1}^{M} \sum_{k'} \hat{\mathcal{F}}_{k'}^{(r)*} \mathbf{E}_{k'}(\boldsymbol{\theta})^{-1} \hat{\mathcal{F}}_{k'}^{(r)} \tag{22}$$

and is identical to (17).

5 Discussion

We now present a discussion of the Bayesian PSD method from both theoretical and computational point of view. The discussion assumes that a single set of consecutive time history of ambient data is given, as is typically the case.

5.1 Underlying assumptions

The PSD method involves statistical proxy and averaging concepts for making inference as it uses as data the sample PSD matrix averaged over independent segments. The FFT method shows that these are not necessary as one can make Bayesian inference directly based on the FFT of the whole set of measured time history. In addition to those modeling assumptions of the FFT method, the PSD method assumes that the segments are independent. When $M = 1$ the NLLF (14) of the PSD method is mathematically identical to that of the FFT method in

(6), but in this case it is no longer the correct NLLF of the PSD method (unless $n=1$) because it is only valid for $M \geq n$.

5.2 Deviation from long data asymptotics

The apparent equivalence shown in Section 4 can deviate significantly from reality because the data length of each segment (T_d / M) need not be large even if T_d is large. This weakens the validity of the long segment assumption in the PSD method and leads to modeling error. The latter results from discrepancy between the theoretical distribution of the sample PSD (assuming long data length) and the actual distribution. In this regard, the validity of three theoretical results is critical to the likelihood function and hence the NLLF in (14):

1. the independence of the sample PSD matrices at different frequencies, which allows the likelihood function to be written simply as a product of the PDFs at different frequencies;
2. the analytical form of (11), i.e., the complex Wishart distribution;
3. the analytical expression in (7) for the expectation of the sample PSD matrix, which facilitates mathematical analysis and computation of the NLLF so that fast algorithms can be derived.

The effect of the segment length on these aspects shall be illustrated later through a numerical study in Section 6. The results reveal that the first two aspects are robust to the segment length but the third can be far from reality when the number of segments is not small, leading to bias in the identification results.

It should be noted that when the PSD method was first formulated in [10] the exact theoretical expression for $E[\hat{\mathbf{E}}_{k'} | \boldsymbol{\theta}]$ in terms of the modal parameters was used, which was valid regardless of the segment length. Thus, only the first two aspects above are relevant to the original formulation. The expression involves the FFT of the theoretical correlation function, or equivalently the integral of the theoretical PSD matrix with the Fejer kernel [25][26], and is not conducive to mathematical analysis or computations. The discussion here refers to the efficient form of the PSD method that uses the asymptotic form of $E[\hat{\mathbf{E}}_{k'} | \boldsymbol{\theta}]$ in (7), for which the third aspect above is also relevant.

5.3 Computational aspects

Since the NLLF of the PSD method shares the same mathematical form as the NLLF of the FFT approach, the algorithms for efficient computation of posterior statistics in principle can be used interchangeably. In fact, the derivation of the algorithm developed in [22] was similar to that in [15]. Regardless of the algorithm developed for computing the posterior statistics, the NLLF in the formulation should be strictly obeyed, for otherwise the Bayesian nature of the method will be lost. That is, the MPV should minimize the NLLF and the covariance matrix should be equal to the inverse of the Hessian of the NLLF at the MPV. In this regard, the algorithm in [22] is not numerically consistent. It obtains the MPV of the 'spectral parameters', i.e., those other than the mode shapes, and the mode shapes separately in two stages. The MPV of the spectral parameters minimizes a surrogate NLLF formulated based on the trace of the sample PSD matrix as data rather than the sample PSD matrix itself. In addition, the surrogate NLLF assumes that the number of segments M is large so that the trace of the sample PSD matrix is asymptotically Gaussian. Unless the original time history data is 'super-long' so that both T_d / M and M are large, the large M assumption contradicts with the long segment assumption.

6 Illustrative example

In this section we present a numerical study with synthetic data to illustrate the issues discussed in the last section. We shall first illustrate the effect of the number of segments on the distribution of the sample PSD and its correlation with different frequencies. A similar study is also applied to the sample PSD averaged over different segments, which is used as the data for modal identification in the Bayesian PSD method. This leads naturally to the effect on the modal identification results, where the potential bias in terms of the posterior MPV and variance shall be investigated. As we shall see, as the number of segments increases and hence the segment length is shortened, the distribution of the sample PSD deviates from the asymptotic value assumed in the identification theory. This can lead to a significant bias (over-estimation) in the identified damping values.

6.1 Synthetic data

Consider a time history of SDOF (single degree of freedom) synthetic acceleration measurement $\{\hat{\ddot{x}}_j\}_{j=1}^N$ of duration $T_d = 1000$ sec and sampling frequency 100 Hz, i.e., $N = 1000 \times 100 = 100{,}000$. The data is modeled as

$$\hat{\ddot{x}}_j = \ddot{x}(t_j) + \varepsilon_j \qquad j = 1,..., N \qquad (23)$$

where $\{\varepsilon_j\}_{j=1}^N$ is i.i.d. (independent and identically distributed) white noise with a PSD of $S_e = 1(\mu g)^2 / Hz$ (typical of force-balanced accelerometers); $\ddot{x}(t_j)$ is the acceleration response of an SDOF structure satisfying

$$\ddot{x}(t) + 2\zeta\omega\dot{x}(t) + \omega^2 x(t) = p(t) \qquad (24)$$

with $\omega = 2\pi f$ (rad/sec), $f = 1$Hz (natural frequency) and $\zeta = 1\%$ (damping ratio); $p(t)$ is white noise excitation with a PSD of $S = 0.04(\mu g)^2$/Hz. Figure 1 shows the sample PSD calculated from a typical set of data. The horizontal bar in the figure indicates the selected frequency band for modal identification in Section 6.5, which can be ignored for the moment.

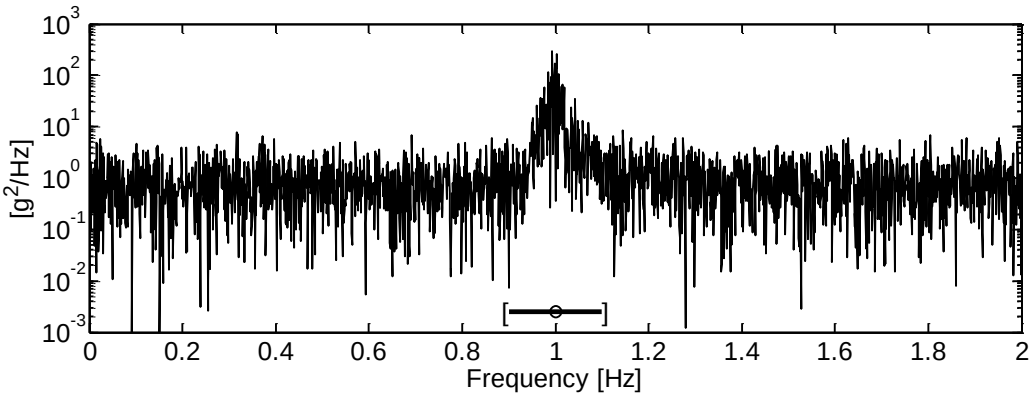

Figure 1 Sample PSD calculated from a typical set of data of duration $T_d = 1000$ sec. Horizontal bar indicates the selected frequency band for modal identification

6.2 Frequency resolution

In the context of the PSD method, suppose we divide the original data set into M non-overlapping segments, so that each has a duration of T_d / M. The frequency interval, i.e., the

frequency difference between neighboring FFT ordinates, is equal to $\Delta f = M / T_d = M / 1000$ (Hz). Figure 2 shows the reduction of frequency resolution with increasing M, overlaying the dynamic amplification factor for illustration.

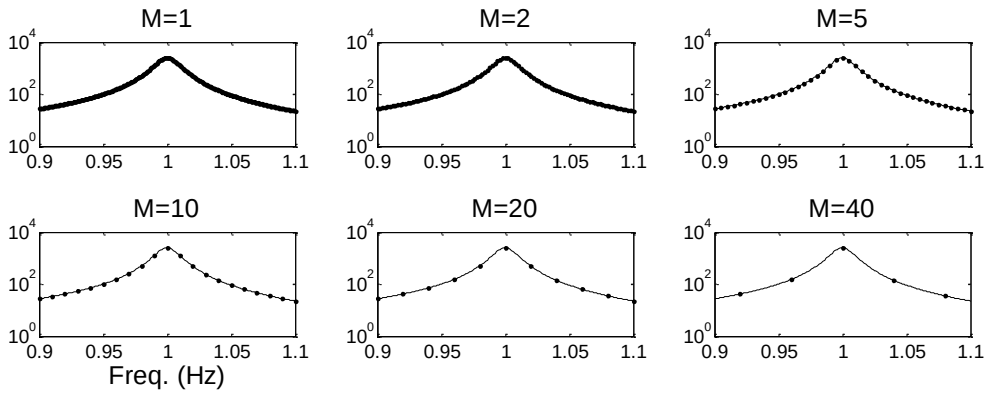

Figure 2 Frequency resolution for different number of segments M. Solid line – dynamic amplification factor. Dots – FFT ordinates

6.3 Sample PSD

We next illustrate the effect of shortened data segment on the distribution of the sample PSD. Let $\hat{S}_{k'}$ be the sample PSD calculated from a particular segment (of length T_d / M), say, the first segment, at a particular frequency $\mathrm{f}_{k'}$. That is,

$$\hat{S}_{k'} = | \hat{F}_{k'}^{(1)} |^2 \tag{25}$$

where, according to (9),

$$\hat{F}_{k'}^{(1)} = \sqrt{\frac{2M\Delta t}{N}} \sum_{j=1}^{N/M} \hat{\ddot{x}}_j^{(1)} e^{-2\pi \mathrm{i} M (j-1)(k'-1)/N} \tag{26}$$

The theoretical PDF of $\hat{S}_{k'}$ is given by (11) with $\hat{\mathbf{E}}_{k'} \equiv \hat{S}_{k'}$, $n = 1$ and $M = 1$. This gives an exponential PDF

$$p(\hat{S}_{k'}) = \frac{1}{E[\hat{S}_{k'}]} \exp\left\{ - \frac{\hat{S}_{k'}}{E[\hat{S}_{k'}]} \right\} \tag{27}$$

where $E[\hat{S}_{k'}]$ is the expectation of $\hat{S}_{k'}$. If the duration of the segment (T_d / M) is sufficient long (theoretically infinite), $E[\hat{S}_{k'}]$ is given by the asymptotic expression in (7) with $n=1$, $\mathbf{\Phi}=1$ and $\mathbf{H}_k = SD_k$:

$$E[\hat{S}_{k'}] = SD_k + S_e \qquad \text{(assuming } T_d / M \text{ large)} \qquad (28)$$

where

$$D_k = [(\beta_k^2 - 1)^2 + (2\zeta\beta_k)^2] \qquad (29)$$

is the dynamic amplification factor and $\beta_k = f / \mathrm{f}_k$; f_k is the FFT frequency.

6.3.1 Expectation

For illustration consider the sample PSD $\hat{S}_{k'}$ at the natural frequency, i.e., $\mathrm{f}_{k'} = 1\,\mathrm{Hz}$. Figure 3 compares the asymptotic value of $E[\hat{S}_{k'}]$ from (28) with the actual value estimated from 100,000 independent trials. These two values almost coincide at $M=1$, which shows that the original data length $T_d = 1000\,\mathrm{sec}$ is sufficiently long for a good asymptotic long data length approximation. When M increases, the actual value is less than the asymptotic value. This reduction is often referred as 'spectral leakage' in signal processing. Thus, when M is not small the duration of each segment T_d / M is no longer large even when T_d is.

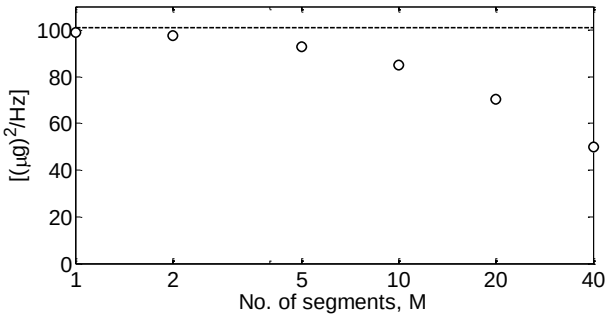

Figure 3 Expectation of $\hat{S}_{k'}$ of segment with duration T_d / M. Dashed line – asymptotic formula in (28); circle – sample estimate from 100,000 trials. Discrepancy between circle and dashed line reflects spectral leakage. Note that the x-axis is plotted on log scale.

6.3.2 Probability distribution

Figure 4 shows the actual PDF of $\hat{S}_{k'}$ (solid line) estimated from the same 100,000 independent trials. The theoretical PDF in (27) with $E[\hat{S}_{k'}]$ given by (28) (which assumes large T_d / M) is also shown with dashed line in the figure for comparison. When $M = 1$, the actual PDF agrees well with the asymptotic PDF. As M increases their difference increases. When $M \geq 20$ (say) the difference is visually significant and the theoretical PDF fails to give a reasonable approximation to the actual PDF. As discussed in Section 5.2, the departure of the theoretical PDF from reality can be attributed to the departure of its mathematical form (complex Wishart distribution) and the departure of the asymptotic value of $E[\hat{S}_{k'}]$ in (28) from the actual value. Figure 5 shows the results analogous to Figure 4 except that theoretical value (dashed line) is now calculated with $E[\hat{S}_{k'}]$ estimated from the sample mean of the 100,000 trials. The agreement between the solid and dashed line reveals that discrepancy between the PDFs observed in Figure 4 is mainly attributed to the departure of (28) from reality.

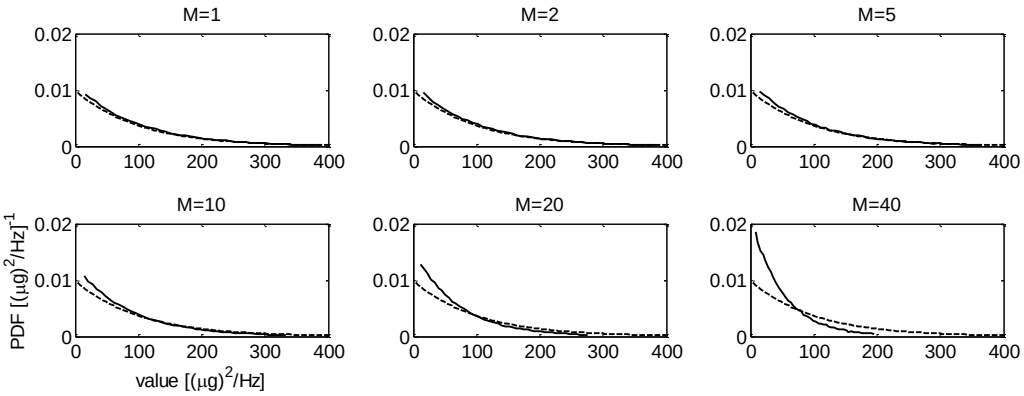

Figure 4 PDF of sample PSD based on data length T_d / M ($T_d = 1000$ sec). Solid line – sample estimate from 100,000 trials; dashed line – (27) with $E[\hat{S}_{k'}]$ given by (28)

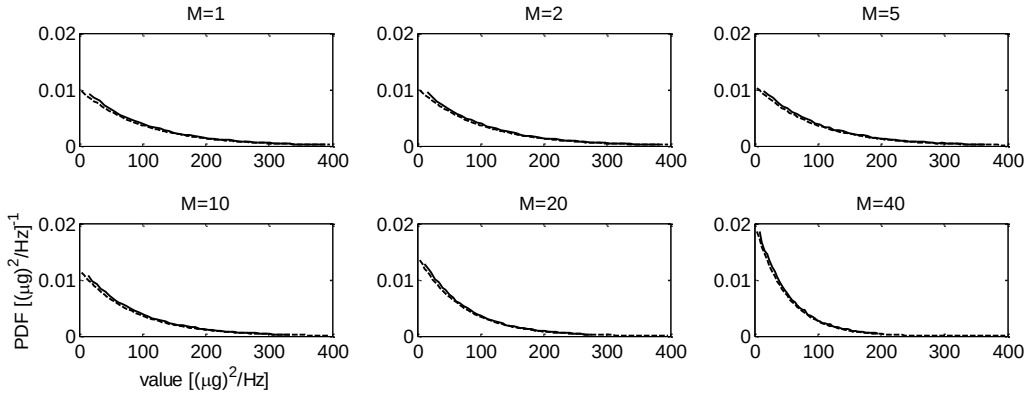

Figure 5 PDF of sample PSD based on data length T_d / M ($T_d = 1000$ sec). Solid line – sample estimate from 100,000 trials; dashed line – (27) with $E[\hat{S}_{k'}]$ estimated from sample mean of 100,000 trials

6.3.3 Correlation among different frequencies

Figure 6 shows the correlation coefficient of the sample PSD at the natural frequency (1Hz) with other neighboring FFT frequencies, estimated from the same 100,000 samples used in the previous sections. It is seen that the correlation is practically zero regardless of M. It is therefore valid to assume that the sample PSDs at different frequencies are independent even when M is not small. Thus, the effect of the number of segments on the independence of FFT at different frequencies can be considered insignificant.

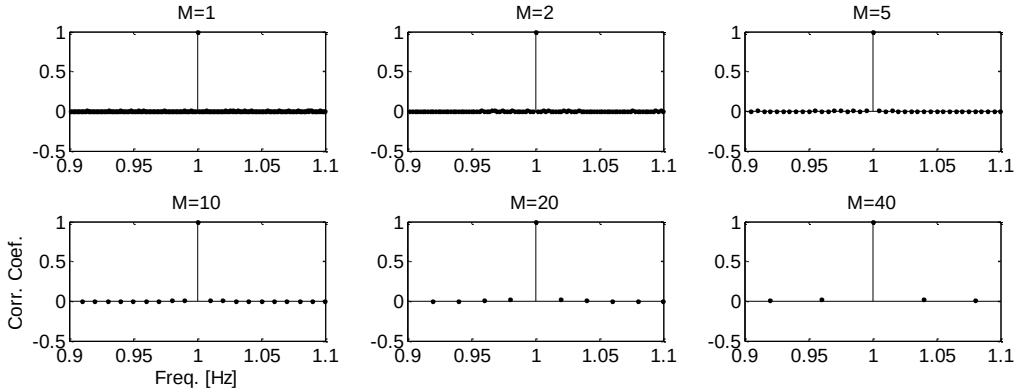

Figure 6 Correlation coefficient of sample PSD at the natural frequency (1Hz) with other frequencies, estimated from 100,000 trials

6.4 Sample PSD averaged over segments

Consider now the sample PDF averaged over the M segments, i.e., analogous to (10),

$$\hat{E}_{k'} = \frac{1}{M}\sum_{r=1}^{M} \hat{S}_{k'}^{(r)} \tag{30}$$

where $\hat{S}_{k'}^{(r)}$ denotes the sample PSD of the r-th data segment. Note that $\hat{S}_{k'}^{(1)}$ was abbreviated to $\hat{S}_{k'}$ in Section 6.3 for simplicity in notation. Clearly, $E[\hat{E}_{k'}] = E[\hat{S}_{k'}]$. Using (11) with $n = 1$, the PDF of $\hat{E}_{k'}$ is given by

$$p(\hat{E}_{k'}) = \frac{M^M \hat{E}_{k'}^{M-1}}{(M-1)!E[\hat{S}_{k'}]^M} \exp\left\{-M\frac{\hat{E}_{k'}}{E[\hat{S}_{k'}]}\right\} \qquad (M \geq 1) \tag{31}$$

6.4.1 Probability distribution

Figure 7 shows the actual PDF of $\hat{E}_{k'}$ (solid line) estimated from the same 100,000 independent trials in Section 6.3. The theoretical PDF in (31) with $E[\hat{S}_{k'}]$ given by (28) (which assumes large T_d / M) is also shown with dashed line in the figure for comparison. When $M = 1$, the actual PDF agrees well with the asymptotic PDF. Their discrepancy increases with M and is seen to be visually significant when $M \geq 10$. This discrepancy is responsible for the bias in the modal identification results presented in Section 6.5 later.

Figure 8 show the results analogous to Figure 7 except that theoretical value (dashed line) is now calculated with $E[\hat{S}_{k'}]$ estimated from the sample mean of the 100,000 trials. Again, the agreement between the solid and dashed line reveals that the departure of (28) from reality is the main contributor to the discrepancy seen in Figure 7.

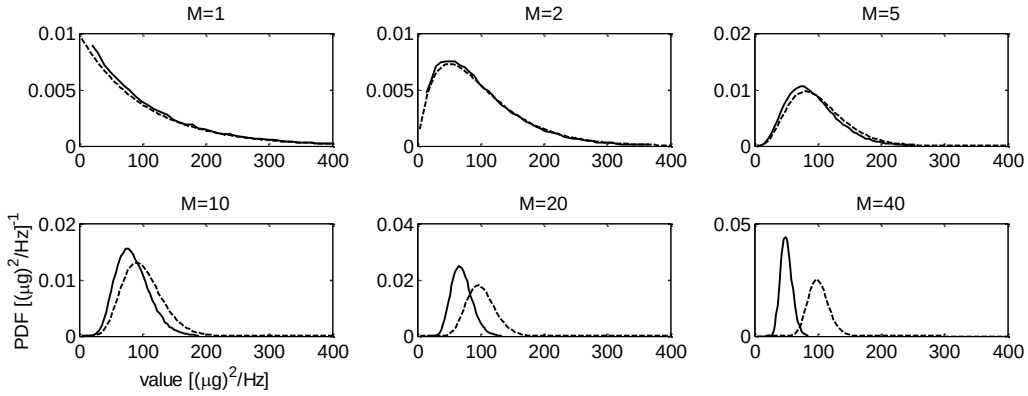

Figure 7 PDF of sample PSD averaged over M segments, each of length T_d / M ($T_d = 1000$ sec). Solid line – sample estimate from 100,000 trials; dashed line – (31) with $E[\hat{S}_{k'}]$ given by (28)

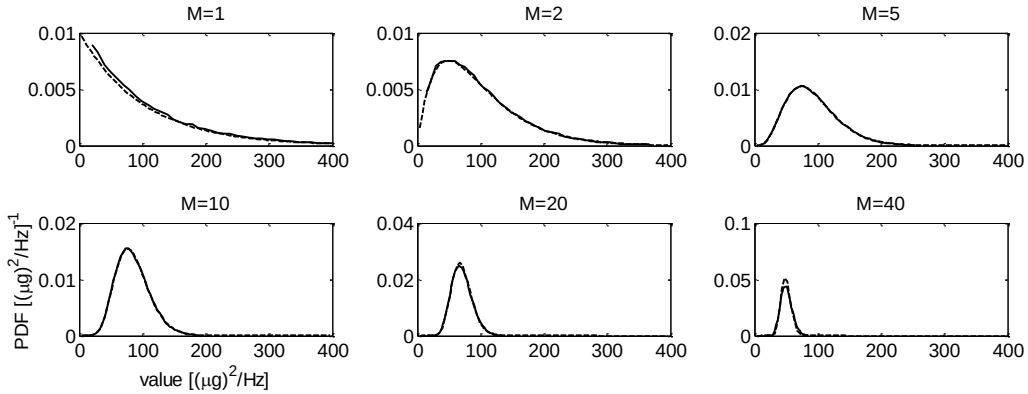

Figure 8 PDF of sample PSD averaged over M segments, each of length T_d / M ($T_d = 1000$ sec). Solid line – sample estimate from 100,000 trials; dashed line – (31) with $E[\hat{S}_{k'}]$ estimated from sample mean of the same 100,000 trials

6.4.2 Correlation among different frequencies

Similar to Figure 6, Figure 9 shows the correlation of the PSD at the natural frequency with other FFT frequencies. The correlation is generally higher compared to Figure 6 but is still small. This reveals that the independence at different frequencies is fairly robust to the segment length and it can be a good approximation even when the number of segments is not small.

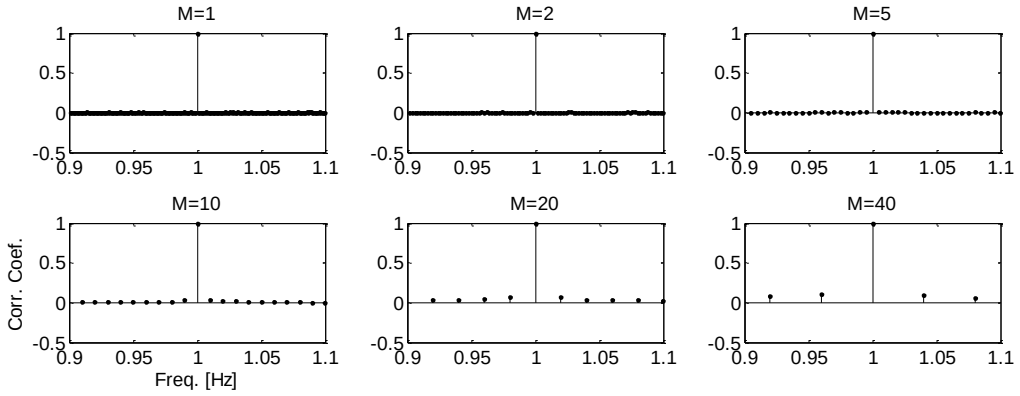

Figure 9 Correlation coefficient of averaged PSD at 1Hz with other frequencies, estimated from 100,000 trials

6.5 Modal identification

We next investigate the modal identification results using the PSD method based on the likelihood function that assumes large segment length (T_d / M), which is currently the only setting that allows fast algorithm to be developed and the method be practically implemented. As we shall see, as the number of segments increases the identified damping ratio is increasingly biased high.

Consider modal identification using the PSD approach with $T_d = 1000$ sec of ambient data generated in the last section, where the PSD calculated from a typical set of data was shown in Figure 1. As indicated in the figure, the spectral information within the frequency band $[0.9,1.1]$ Hz is used for modal identification. To provide a perspective on the quality and amount of the data considered here, the modal signal-to-noise (s/n) ratio is [14]

$$\gamma = \frac{S}{4S_e\zeta^2} = \frac{0.04}{4(1)(0.01)^2} = 100 \qquad (32)$$

This is typical in field tests under normal wind conditions and modes with this s/n ratio can be easily identified. On the other hand, a data length of $T_d = 1000$ sec is equivalent to $N_c = T_d f = (1000)(1)=1000$ natural periods, which is sufficient for identifying the damping ratio with good accuracy [27][28].

To investigate the quality of the identification results we generate 1000 independent sets of data and perform modal identification for each set. For each set of data we perform modal identification with different number of segments M. The discussion here focuses on the posterior MPV and posterior c.o.v. (coefficient of variation = posterior standard deviation/posterior MPV) of the natural frequency and damping ratio.

6.5.1 Posterior MPV

Figure 10 shows the posterior MPVs (dots) of the natural frequency and damping ratio of 1000 trials of data. The square shows the sample mean of the MPVs and the dashed line shows the exact value used for generating the data for identification. Note that the results for $M=1$ is identical to those one could have obtained using the Bayesian FFT method. For the natural frequency, the results are qualitatively the same for different M. For the damping ratio, however, as M increases the MPVs are increasingly biased high. This bias can be explained by noting that the likelihood function assumes the asymptotic distribution for the sample PSD but it is biased high (Figure 7). Intuitively, this drives the MPV of the damping ratio in the identification process (minimizing the NLLF) to higher values so as to drive the theoretical PDF of the PSD to the low side to 'match' probabilistically the sample PSD calculated from the data. The fact that the results are practically unbiased when $M=1$ reflects that the total data duration of $T_d=1000\,\text{sec}$ is indeed sufficiently long for the Bayesian FFT method to be valid, but the same is not true for the PSD method when the number of segments is not small.

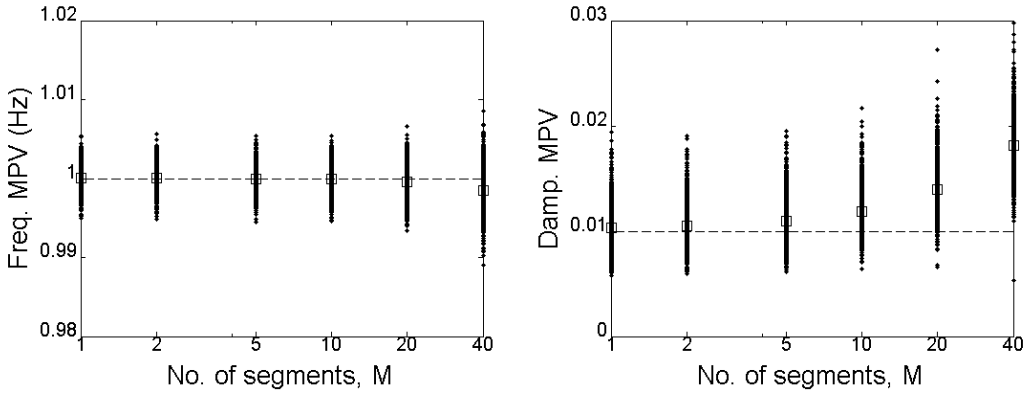

Figure 10 Posterior MPVs of modal properties obtained from Bayesian PSD method with different number of segments (*M*). Dots – results from 1000 trials. Square – sample mean of

1000 trials. Dashed line – exact value that generated the data. Note that the horizontal axis is plotted on log-scale. Results show that as *M* increases the damping ratio identified by the PSD approach is biased high.

6.5.2 Posterior c.o.v.

We next investigate the posterior c.o.v. of the natural frequency and damping ratio. Figure 11 shows the posterior c.o.v.s of the 1000 trials (dots). The square shows the root mean square (RMS) value of the c.o.v.s. The RMS value is considered rather than simply the sample mean because it is the proper form of averaging for variability. The solid line shows the sample c.o.v. of the MPVs shown in Figure 10. There is a slight increase in the posterior c.o.v.s and sample c.o.v.s as the number of segments increases. Nevertheless, the RMS value of the posterior c.o.v. agree well with the sample c.o.v. of MPV, showing that it properly reflects the posterior variability of the identified values around their sample mean (which is nevertheless biased for the damping ratio).

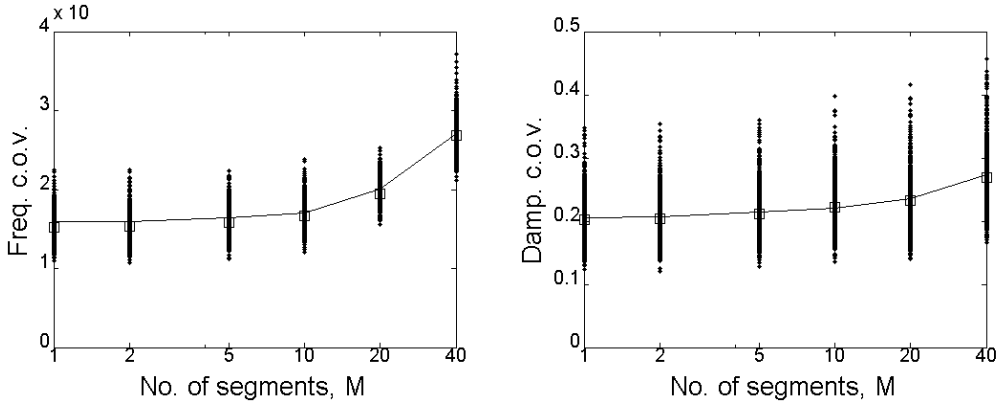

Figure 11 Posterior c.o.v. of modal properties obtained from Bayesian PSD method with different number of segments (*M*). Dots – results from 1000 trials. Square – sample mean of 1000 trials. Solid line – sample c.o.v. of MPVs. Note that the horizontal axis is plotted on log-scale. Results show that the posterior c.o.v. is generally consistent with the sample c.o.v. of MPVs.

It was shown previously [29] that if there is no modeling error the Bayesian measure and frequentist measure of identification uncertainty should agree. In the present case the Bayesian measure is the RMS value of the posterior c.o.v. and the frequentist measure is the

sample c.o.v. of the MPVs. Despite the agreement between the square and solid line in Figure 11, the Bayesian and frequentist measure do not agree for the damping ratio because the sample c.o.v. of the MPVs were calculated based on the (biased) sample mean rather than the actual damping value that generated the data. That is, had the sample variance (from which the c.o.v. is calculated) of the MPVs been calculated (correctly) as the average squared difference from the exact value of damping ratio (1%) rather than the sample mean, the resulting sample c.o.v. (frequentist measure) would not have agreed with the RMS of posterior c.o.v. (Bayesian measure).

7 Conclusions

Averaging concepts are useful for producing a smooth spectrum for visualization but as far as modal identification is concerned it is fundamentally unnecessary, as evidenced by the Bayesian FFT formulation. Despite the apparent mathematical equivalence in the likelihood function of the Bayesian FFT and PSD method, the FFT in the latter is calculated based on shorter data segments and so is more vulnerable to deviation from long data asymptotics. This can lead to modeling error and bias in the identification results if the likelihood function is based on long data asymptotics. A single time history can be considered sufficiently long for the Bayesian FFT method but the same need not be true for the PSD method, depending on the number of segments divided for averaging.

Long data asymptotics is responsible for the independence of FFTs at different frequencies and the asymptotic form of the likelihood function. The numerical study reveals that the independence of FFTs is insensitive to the number of segments but the same is not true for the asymptotic form of the likelihood function, in particular, the theoretical expectation of the sample PSD matrix. When the number of segments is not small the actual expectation of the sample PSD is significantly smaller than the asymptotic value. The resulting bias is therefore essentially due to spectral leakage. The numerical study reveals that the bias in the damping ratio is insignificant when the FFT is calculated with a data length of 500 natural periods. This applies to both the Bayesian FFT and PSD method. Implementing the PSD method with a given set of data, the number of segments should be as small as possible. As the complex Wishart distribution (basis of the PSD method) is valid as long as the number of segments is

greater than the number of measured DOFs, the number of segments in the PSD method should be simply taken as the number of measured DOFs.

Although for simplicity of discussion the numerical study considers only one measured DOF, the findings are generally applicable to multiple measured DOFs. This is supported by recently derived uncertainty laws [27][28], where the measured DOFs affect only the modal s/n ratio but which has a diminishing effect for good quality data. The findings are also applicable to close modes as long as the modal s/n ratio is reasonable (e.g., >100) and the measured mode shapes are linearly independent, as evidenced by a recent numerical study [30]. Guidelines on non-Bayesian identification methods based on PSD may be referred to [31].

8 Acknowledgments

This work in this paper is supported by University of Liverpool grant EGG10034.

9 References

[1] H. Wenzel, D. Pichler, Ambient Vibration Monitoring, John Wiley & Sons, UK, 2005.

[2] J.M.W. Brownjohn, Ambient vibration studies for system identification of tall buildings, Earthquake Engineering and Structural Dynamics, 32 (2003) 71-95.

[3] B. Peeters, G. De Roeck, Stochastic system identification for operational modal analysis: a review, Journal of Dynamical Systems, Measurement and Control, 123 (2001) 659-667.

[4] R. Brincker, L. Zhang, P. Anderson, Modal identification of output-only systems using frequency domain decomposition, Smart Materials and Structures, 10 (3) (2001) 441-455.

[5] E. Reynders, R. Pintelon, G. De Roeck, Uncertainty bounds on modal parameters obtained from stochastic subspace identification, Mechanical Systems and Signal Processing, 22, 948-969, 2007.

[6] R.T. Cox, The algebra of probable inference, Johns Hopkins Press, Baltimore, 1961.

[7] E.T. Jaynes, Probability Theory: The logic of Science. Cambridge University Press, 2003.

[8] J.L. Beck, Bayesian system identification based on probability logic, Structural Control and Health Monitoring, 17 (7) (2010) 825-847.

[9] K.V. Yuen, L.S. Katafygiotis, Bayesian time-domain approach for modal updating using ambient data, Probabilistic Engineering Mechanics, 16 (3) (2001) 219-231.

[10] K.V. Yuen, L.S. Katafygiotis, Bayesian spectral density approach for modal updating using ambient data, Earthquake Engineering and Structural Dynamics, 30 (2001) 1103-1123.

[11] K.V. Yuen, L.S. Katafygiotis, Bayesian Fast Fourier Transform approach for modal updating using ambient data, Advances in Structural Engineering, 6 (2) (2003) 81-95.

[12] S.K. Au, F.L. Zhang, Y.C. Ni, Bayesian operational modal analysis: theory, computation, practice, Computers and Structures, 126 (2013) 3-14.

[13] J.L. Beck, L.S. Katafygiotis, Updating models and their uncertainties, I: Bayesian statistical framework, Journal of Engineering Mechanics, ASCE, 124 (4) (1998) 455-461.

[14] S.K. Au, Fast Bayesian FFT method for ambient modal identification with separated modes, Journal of Engineering Mechanics, ASCE, 137 (3) (2011) 214-226.

[15] S.K. Au, Fast Bayesian ambient modal identification in the frequency domain, Part I: posterior most probable value, Mechanical Systems and Signal Processing, 26 (1) (2012) 60-75.

[16] S.K. Au, Fast Bayesian ambient modal identification in the frequency domain, Part II: posterior uncertainty, Mechanical Systems and Signal Processing, 26 (1) (2012) 76-90.

[17] S.K. Au, F.L. Zhang, Fast Bayesian ambient modal identification incorporating multiple setups, Journal of Engineering Mechanics, 138 (7) (2012) 800-815.

[18] F.L. Zhang, S.K. Au, H.F. Lam, Assessing uncertainty in operational modal analysis incorporating multiple setups using a Bayesian approach, Structural Control and Health Monitoring, 22(3), 395-416, 2015.

[19] S.K. Au, P. To, Full-scale validation of dynamic wind load on a super-tall building under strong wind, Journal of Structural Engineering, ASCE, 138 (9) (2012) 1161-1172.

[20] S.K. Au, Y.C. Ni, F.L. Zhang, H.F. Lam, Full scale dynamic testing of a coupled slab system, Engineering Structures, 37 (2012) 167-178.

[21] S.K. Au, F.L. Zhang, P. To, Field observations on modal properties of two tall buildings under strong wind, Journal of Wind Engineering and Industrial Aerodynamics, 101 (2012) 12-23.

[22] J.Y. Wang, L.S. Katafygiotis, A two-stage fast Bayesian spectral density approach for ambient modal analysis, Part I: Posterior most probable value and uncertainty, Mechanical Systems and Signal Processing, 54-55, 139-155.

[23] N.R. Goodman, Statistical analysis based on a certain multivariate complex Gaussian distribution, Annals of Mathematical Statistics, 34, 152–177, 1963.

[24] M.S. Srivastava, On the Complex Wishart Distribution, Annals of Mathematical Statistics, 36(1), 313-315, 1965.

[25] D.R. Brillinger, Time series: Data analysis and theory, Holden-Day, Inc., San Francisco, 1981.

[26] K.V. Yuen, L.S. Katafygiotis, Spectral density estimation of stochastic vector processes, Probabilistic Engineering Mechanics, 17 (2002) 265-272.

[27] S.K. Au, Uncertainty law in ambient modal identification, Part I: theory, Mechanical Systems and Signal Processing, 48 (1-2) (2014) 15-33.

[28] S.K. Au, Uncertainty law in ambient modal identification, Part II: implication and field verification, Mechanical Systems and Signal Processing, 48 (1-2) (2014) 34-48.

[29] S.K. Au, Connecting Bayesian and frequentist quantification of parameter uncertainty in system Identification, Mechanical Systems and Signal Processing, 29 (2012) 328-342.

[30] Y.C. Zhu, S.K. Au, S. Jones, Identification uncertainty of close modes in operational modal analysis, In Proceedings of 12th International Conference on Applications of Statistics and Probability in Civil Engineering, 12-15 July 2015, Vancouver, Canada.

[31] J.S. Bendat, A.G. Piersol, Random data: Analysis and Measurement Procedures, 4th Edition, Wiley, New York, 2010.