Uncertainty Law in Ambient Modal Identification

Part II: Implication and Field Verification

Siu-Kui Au*

Center for Engineering Dynamics and Institute for Risk and Uncertainty

University of Liverpool

Abstract

This paper presents a qualitative analysis of the uncertainty laws for the modal parameters identified in a Bayesian approach using ambient vibration data, based on the theory developed in the companion paper. The uncertainty laws are also appraised using field test data. The paper intends to provide insights for planning ambient vibration tests and managing the uncertainties of the identified modal parameters. Some typical questions that shall be addressed are: To estimate the damping ratio to within 30% of posterior coefficient of variation (c.o.v), what is the minimum data duration? Will deploying an additional accelerometer significantly improve the accuracy in damping (or frequency)? Answers to these questions based on this work can be found in the Conclusions. As the Bayesian approach allows full use of information in the data for given modeling assumptions, the uncertainty laws obtained in this work represent the lower limit of uncertainty (estimation error) that can be achieved by any method (Bayesian or non-Bayesian).

Keywords

Ambient vibration; operational modal analysis; spectral analysis; uncertainty law

* Corresponding author. E-mail: siukuiau@liverpool.ac.uk. Office phone: +44 (0)151 794 5217. Formerly City University of Hong Kong.

1. Introduction

Uncertainty in the identified modal parameters is an important aspect to manage in planning an ambient vibration test. Channel noise, sampling rate and bandwidth of data can be well-controlled these days, thanks to advances in modern sensing and data acquisition technology. The attributes that often need to be decided on a case-by-case basis include, among others, the number of sensors, the location of sensors and the data duration. Mechanical concepts, together with experience with sensing and data acquisition hardware, can help configure these attributes. The number of sensors is constrained by availability and budget. The location of sensors depends on the mode shapes expected to be found. Logistics and accessibility constraints are critical factors, although progress in theoretical development cannot be overlooked [1][2][3][4]. The duration of data is often decided by rule of thumb, e.g., 1000 natural periods of the lowest mode of interest. It can be constrained by the available time, e.g., on a construction site. In principle, increasing the data length is expected to improve accuracy by virtue of increasing the amount of information. Identifying parameters using an extended data length, however, can increase modeling error risk [5]. For example, assuming a time-invariant model, the damping ratio identified based on a long period of data where the response amplitude has changed significantly can at best represent the average value of the actual amplitude-dependent damping [6][7][8][9]. The damping ratio is an important parameter in applications as it directly affects the magnitude of dynamic response. However, there is no commonly accepted method for reliable prediction at the design stage. It is also difficult to estimate from measured data, due to, e.g., measurement error, modeling error and amplitude-dependence. Methods that rely on statistical proxies (e.g., averaging) are vulnerable to bias [10][11][12][13]. It is necessary to quantify the uncertainty associated with damping estimates so that the results can be interpreted in the right context.

A Bayesian FFT approach allows full extraction of information contained in the data for modal identification [14]. The raw FFTs instead of their averaged counterparts are used for statistical inference, therefore eliminating possible distortion due to averaging or other signal processing artifacts. Based on the same data and modeling assumptions, no non-

Bayesian method can be more informative about the modal parameters than the Bayesian method. The uncertainty laws therefore represent the lower limit of estimation error that can be achieved any method (Bayesian or non-Bayesian) for given data and modeling assumptions.

An asymptotic analysis of the 'posterior uncertainty' (i.e., given data) of modal parameters in a Bayesian context has been performed in the companion paper. Assuming well-separated modes, small damping and sufficient amount of data, asymptotic expressions for the posterior covariance matrix of modal parameters have been derived. The results are *remarkably simple*. This paper presents a qualitative analysis of the uncertainty laws to yield insights for planning ambient vibration tests and managing the uncertainties of the identified modal properties. The uncertainty laws are also verified using field test data.

2. Qualitative analysis

We first recall the main results derived in the companion paper. To the leading order, the (squared) posterior coefficient of variation (c.o.v.=standard deviation/most probable value) of the natural frequency f, damping ratio ζ, PSD (power spectral density) of modal force S, and PSD of prediction error S_e, are given by

$$\delta_f^2 \sim \frac{\zeta}{2\pi N_c B_f(\kappa)},\ \delta_\zeta^2 \sim \frac{1}{2\pi\zeta N_c B_\zeta(\kappa)},\ \delta_S^2 \sim \frac{1}{N_f B_S(\kappa)},\ \delta_{S_e}^2 \sim \frac{1}{(n-1)N_f} \tag{1}$$

where n is the number of measured degrees-of-freeom (dofs); $N_c = T_d / T$ ($T_d =$ data duration; $T =$ natural period) is the data length as a multiple of the natural period; $N_f = 2\kappa\zeta N_c$ is the number of frequency ordinates in the selected band $f(1 \pm \kappa\zeta)$ and κ is the 'bandwidth factor';

$$B_f(\kappa) = \frac{2}{\pi}(\tan^{-1}\kappa - \frac{\kappa}{\kappa^2+1}),\ B_\zeta(\kappa) = \frac{2}{\pi}\left[\tan^{-1}\kappa + \frac{\kappa}{\kappa^2+1} - \frac{2(\tan^{-1}\kappa)^2}{\kappa}\right]$$

$$B_S(\kappa) = 1 - 2(\tan^{-1}\kappa)^2\kappa^{-1}(\tan^{-1}\kappa + \frac{\kappa}{\kappa^2+1})^{-1} \tag{2}$$

are 'data length factors' that depend only on κ.

The bandwidth factor κ is a dimensionless parameter that depends on the frequency band selected by the user, which must trade off between modeling error and the information included for modal identification. Theoretically, the wider the selected band (hence larger κ) the more information for identification. However, widening the band makes the identification model more vulnerable to modeling error regarding single mode and constant PSD of modal force/prediction error within the band.

The posterior covariance matrix of the mode shape $\mathbf{\Phi} \in R^n$ (with normalization $\| \mathbf{\Phi} \|^2 = \mathbf{\Phi}^T \mathbf{\Phi} = 1$) is given by

$$\mathbf{C}_\Phi \sim \frac{\nu\zeta}{N_c B_\Phi(\kappa)} (\mathbf{I}_n - \mathbf{\Phi}\mathbf{\Phi}^T) \tag{3}$$

where $\nu = S_e / S$ is the 'noise-to-environment (n/e) ratio'; and

$$B_\Phi(\kappa) = \tan^{-1}\kappa \tag{4}$$

is the data length factor for the mode shape. The expected Modal Assurance Criterion (MAC) that quantifies the overall uncertainty of the mode shape [15] is given by

$$\overline{\rho} = (1 + \delta_\Phi^2)^{-1/2} \tag{5}$$

where δ_Φ^2 is the sum of principle variances of $\mathbf{C}_\Phi$ given by

$$\delta_\Phi^2 \sim \frac{(n-1)\nu\zeta}{N_c B_\Phi(\kappa)} \tag{6}$$

2.1. Governing scales

The posterior uncertainties in (1) and (3) depend on the following (dimensionless) scales: ζ, ν, κ, N_c or N_f. The damping ratio ζ is a property of the structure. The n/e ratio ν represents a modal noise-to-signal ratio excluding the effect of dynamic amplification. The 'normalized' data length N_c is related to the maximum amount of information available in the data for inferring the mode of interest. On the other hand, κ and $N_f = 2\kappa\zeta N_c$ are related to the amount of information that can be actually utilized. As mentioned before, the bandwidth factor κ depends on the frequency band selected by the

user, which must trade off between modeling error and the information included for modal identification. For example, one can have a long time history of data so that N_c is large. However, in the neighborhood of the natural frequency other unknown colored noise are contributing, such that only a small bandwidth can be used for identification without significant modeling error, which limits κ and N_f.

2.2. Data length effect

The posterior c.o.v.s δ_f, δ_ζ and δ_Φ are inversely proportional to $N_c^{1/2}$, while δ_S and δ_{S_e} are inversely proportional to $N_f^{1/2}$. This inverse square root law is common in statistical estimation. Two different scales of data length are relevant here because f, ζ and $\mathbf{\Phi}$ are related to signals with a particular period, while S and S_e are related to the background environment. For the former, the amount of information is proportional to the number of natural periods in the data. For the latter, it is simply proportional to the number of frequency ordinates in the selected bandwidth without discrimination.

2.3. Usable bandwidth

The posterior uncertainties in (1) and (3) have been written in a form that isolates the effect of the bandwidth factor κ into the data length factor. The corresponding data length factor is an increasing function of κ from zero to one (see Figure 1 later). The remaining part in the formula represents the lower limit of the posterior uncertainty when the full bandwidth can be utilized for identification. For example, for the natural frequency, $\delta_f^2 \geq \zeta / 2\pi N_c$. In reality one is not able to make use of the full bandwidth (from DC to Nyquist) for identification, rendering κ to be finite. The term $N_c B_f(\kappa)$ thus represents the effective data length (as a multiple of natural periods) that can be utilized for identifying the mode when the effect of κ is taken into account.

Figure 1 shows the variation of the data length factors with κ. Note that $B_\Phi > B_f > B_S > B_\zeta$; B_ζ and B_S almost overlap because the term $\tan^{-1}\kappa + \kappa/(\kappa^2+1)$ in

(2) converges quickly to $\pi/2$ for moderate values of κ (say, $\kappa > 2$). For bandwidths that can be typically utilized in practice, say, $\kappa = 6$, the bandwidth factor is about 80% for the natural frequency and 90% for the mode shape. The value for ζ and S is a bit lower, about 60%. The figure suggests that the accuracy in the mode shape or frequency is typically less sensitive to the bandwidth than ζ or S. This can be explained based on common intuitions. A small neighborhood around the resonance peak is sufficient to pin down the natural frequency because here the location in the spectrum matters most. The spectral amplitude ratios at the resonance peak among different dofs can already give a good estimation of the mode shape. The damping ratio is related to the decay of the response PSD around the resonance peak, and so it requires a larger frequency neighborhood for proper estimation. The PSD of modal force is related to the ratio of the response PSD to the dynamic amplification factor. Widening the band directly increases the number of points for estimation. Note that these are just for intuitive reasoning and should not be confused with the Bayesian theory that yields the posterior uncertainty in a fundamental manner based on modeling assumptions and probability logic.

2.4. Measured dofs

The measured dofs dictate the mode shape $\mathbf{\Phi} = [\Phi_1, ..., \Phi_n]^T$. This directly affects the structure of the posterior covariance matrix of the mode shape, $\mathbf{C}_\Phi$, in (3). More importantly but less trivially, the measured dofs affect the n/e ratio $\nu = S_e / S$ and the modal s/n (signal-to-noise) ratio $\gamma = S / 4 S_e \zeta^2$ (see Section 2.5) through the PSD of modal force S. The n/e ratio is a multiplier in the expression of $\mathbf{C}_\Phi$ in (3) and so it directly affects the uncertainty of the mode shape. The modal s/n ratio affects the uncertainty of modal parameters in a characteristic way, as will be discussed in Section 2.5.

The dependence of S on $\mathbf{\Phi}$ arises from the relationship between the physical and modal response, and the scaling of the mode shape. It can be reasoned that if $\mathbf{\Phi}$ is scaled down (i.e., divided) by a factor then S should be scaled up (i.e., multiplied) by the square of

that factor. Together with the unit norm constraint on $\boldsymbol{\Phi}$, this implies that S is an increasing function of n.

The above can be reasoned as follows. Assuming a single mode, the response at the i-th measured dof is theoretically given by $x_i = \Phi_i \eta$ where η is the modal response. If Φ_i is scaled down by c (say) then $\eta(t)$ must be scaled up by c so that the physical response x_i remains unchanged, i.e., $x_i(t) = (\Phi_i / c)(c\eta)$. To give the modal response $c\eta$ the modal excitation must be scaled up by c, and so its PSD (which is associated with second order statistics) must be scaled up by c^2. On the other hand, as the measured mode shape is normalized with its sum of squares equal to 1, the mode shape value of a particular dof must decrease when the mode shape vector is extended to include the additional measured dof. As a result, the mode shape must be scaled down when n increases and hence S must increase (scaled up).

The rate at which S increases with n depends on the mode shape value of the dof incrementally added to the measured set. As an example, if all measured dofs have the same mode shape value then $\Phi_i = 1/\sqrt{n}$ for all $i = 1,...,n$ (so that $\| \boldsymbol{\Phi} \| = 1$) and $S \propto n$. On the other hand, when the additional dof has a mode shape value of zero, S does not change.

More specifically, it can be shown that the PSD of modal force is proportional to the sum of squares of the mode shape values at the measured dofs. Based on the standard structural dynamics equation, the PSD of modal force is given by

$$S_p = \frac{\boldsymbol{\xi}^T \mathbf{S}_F \boldsymbol{\xi}}{(\boldsymbol{\xi}^T \mathbf{M} \boldsymbol{\xi})^2} \tag{7}$$

where $\mathbf{M}$ is the mass matrix, $\mathbf{S}_F$ is the PSD matrix of the forces applied to the structure and $\boldsymbol{\xi}$ is the 'full' mode shape containing all (possibly an infinite number of) dofs. For a constructed structure, S_p in (7) can hardly be calculated because the quantities involved are rarely accessible or difficult to identify. In reality, only the partial mode shape $\boldsymbol{\Phi}$

rather than the full mode shape $\boldsymbol{\xi}$ is identified. The corresponding identified PSD of modal force S (consistent with the scaling of $\boldsymbol{\Phi}$) is a scaled version of S_p.

We next investigate the scaling between S and S_p. Let $\boldsymbol{\xi} = [\xi_1, \xi_2, ..., \xi_n, ...]^T$. Without loss of generality, suppose the first n dofs of $\boldsymbol{\xi}$ correspond to the measured dofs. Then

$$\Phi_i = \frac{\xi_i}{\sqrt{\sum_{j=1}^{n} \xi_j^2}} \tag{8}$$

so that the norm constraint $\boldsymbol{\Phi}^T\boldsymbol{\Phi} = \sum_{j=1}^{n} \Phi_j^2 = 1$ is satisfied. This means that the full mode shape that is compatible with the scaling of $\boldsymbol{\Phi}$ is given by $\boldsymbol{\xi}' = \boldsymbol{\xi} / c$ where $c = \sqrt{\sum_{j=1}^{n} \xi_j^2}$. Replacing $\boldsymbol{\xi}$ in (7) by $\boldsymbol{\xi}'$, the PSD of modal force S that can be identified from data and that is compatible with the scaling of $\boldsymbol{\Phi}$ is given by

$$S = \frac{\boldsymbol{\xi}'^T \mathbf{S}_F \boldsymbol{\xi}'}{(\boldsymbol{\xi}'^T \mathbf{M} \boldsymbol{\xi}')^2} = \frac{(\boldsymbol{\xi}/c)^T \mathbf{S}_F (\boldsymbol{\xi}/c)}{[(\boldsymbol{\xi}/c)^T \mathbf{M} (\boldsymbol{\xi}/c)]^2} = \frac{\boldsymbol{\xi}^T \mathbf{S}_F \boldsymbol{\xi}}{(\boldsymbol{\xi}^T \mathbf{M} \boldsymbol{\xi})^2} c^2 = S_p \sum_{j=1}^{n} \xi_j^2 \tag{9}$$

This equation shows that S is proportional to the sum of squares of the mode shape values at the measured dofs. Note that the equation only provides a conceptual understanding. It is not useful for computing S because S_p and $\boldsymbol{\xi}$ are not available in reality. Rather, S is directly identified from measured data. The effect of the measured dofs on the modal s/n ratio shall be illustrated in Section 2.5 with synthetic data and in Section 4.3 with field data.

2.5. Signal-to-noise effect

The s/n ratio that is fundamental to the identification of a well-separated mode is given by the ratio of the PSD of ambient response to the PSD of prediction error at the resonance peak [16]:

$$\gamma = \frac{S}{4 S_e \zeta^2} \tag{10}$$

This 'modal s/n ratio' is not a unique property of the data channel. It depends critically on ζ through dynamic amplification. The prediction error PSD S_e comes from the

measurement noise in the data channel and modeling error within the selected band, e.g., due to un-modeled contribution from other modes. The PSD of modal force S reflects the intensity of the ambient excitation near the natural frequency. As explained in Section 2.4 it also depends on the mode shape and always decreases with the number of measured dofs n.

The asymptotic posterior c.o.v.s of f, ζ, S and S_e do not depend on the noise-to-environment (n/e) ratio $\nu = S_e / S$ and so they do not depend on the modal s/n ratio, which can be written as $\gamma = 1/4\nu\zeta^2$. Of course, this statement is only correct up to the leading order and it assumes that ζ is small so that the modal s/n ratio is high. This observation, somewhat counter-intuitive at first glance, suggests that when the modal s/n ratio is sufficiently large, further increasing it (e.g., by reducing S_e using better quality equipment or increasing the number of measured dofs n) has insignificant effect on improving the quality of frequency or damping estimates. This happens because the posterior uncertainty does not come only from the prediction error; it also comes from the unknown modal excitation for which a stochastic model has been assumed. Uncertainty in the latter cannot be eliminated by improving the quality of equipment. The only modal parameter whose posterior uncertainty depends on the modal s/n ratio is the mode shape, which shall be discussed in Section 2.6.

Illustrative example (synthetic data)

Here we present an example with synthetic data to illustrate the effect of the measured dofs on the modal s/n ratio and identification uncertainty. Field data examples shall be given in Section 4.3. Consider the horizontal vibration of a ten-storied shear building with uniform floor mass of 100 tons, interstory stiffness of 177kN/mm and damping ratio of 1% in all modes. The natural frequency of the first three modes are 1Hz, 2.98Hz and 4.89Hz. The mode shape of the first mode increases from the bottom to the top of the building. The structure is subjected to i.i.d. (independent and identically distributed) white noise excitation at all floors, each with a PSD of 4.81 N^2/Hz. Synthetic acceleration data is generated at a sampling rate of 100Hz for a duration of 600sec. The

data is contaminated by i.i.d. channel noise with a PSD of $100(\mu g)^2 / Hz$ ($1\mu = 10^{-6}$, $g = 9.81 m/s^2$). Consider identifying the first mode with an increasing number of measured dofs (each with a uniaxial accelerometer) from $n = 2$ to 10. Three cases of sensor layout sequence are considered. In Case 1, the increasing number of sensors are placed from the top to the bottom, i.e., at the roof and 9/F for $n = 2$; at the roof, 9/F and 8/F for $n = 3$ and so on. In Case 2, the sequence is reversed, i.e., at 1/F and 2/F for $n = 2$; at 1/F, 2/F and 3/F for $n = 3$ and so on. In Case 3, all sensors are placed at the roof, i.e., two sensors at the roof for $n = 2$; three sensors at the roof for $n = 3$ and so on. Note that even if the sensors are all placed on the roof their data are not identical, because of channel noise. Modal identification is based on FFT data on the frequency band [0.94, 1.06] Hz, corresponding to a bandwidth factor of $\kappa = 6$.

Figure 2(a) shows the modal s/n ratio $\gamma = S / 4S_e\zeta^2$ (see (10)) calculated using the MPVs identified from data in each case. As expected, γ increases with n in all cases. In Case 1 (dots), the rate (slope) decreases with n because the mode shape value of the additional dofs (lower floors) is decreasing. An analogous argument explains the increasing rate of γ in Case 2 (circles). In Case 3, γ increases almost linearly with n (deviation due to fluctuations in MPVs), because in this case the mode shape value at all the measured dofs (at the roof) are all the same.

Figure 2(b) shows the corresponding (exact) posterior c.o.v. of the damping ratio δ_ζ [16][17]. The posterior c.o.v.s of other modal parameters (which are less critical) are omitted here to simplify discussion. The dashed line in the figure shows the value predicted by the uncertainty law in (1) calculated using the exact parameter values ($\zeta = 1\%$, $N_c = 600$, $\kappa = 6$). The counterpart values calculated using the MPVs identified from data are close to the dashed line; they are not shown to avoid complication in the figure. For Case 2 (circles), δ_ζ decreases with diminishing rate as n increases. In particular, δ_ζ reduces by 30% as n increases from 2 to 3; and by 10% as n increases from 3 to 4. This diminishing rate of uncertainty reduction with n is intimately

related to the insensitivity of δ_ζ to γ when γ is large. In fact, for Case 1 (dots) and Case 3 (square), δ_ζ is insensitive to n right from the beginning because γ is already 50 when $n = 2$ (see Figure 2(a)). This is further explained in Figure 3, which plots the value of δ_ζ versus γ in all cases. As expected, δ_ζ converges to the uncertainty law (dashed line) as γ increases. All points lie almost on the same curve, indicating that the non-trivial influence of the measured dofs on δ_ζ can be explained essentially by the relationship between δ_ζ and γ.

2.6. Mode shape uncertainty

Since the posterior covariance matrix of the mode shape is not a diagonal matrix, the mode shape values at different measured dofs are correlated with each other. This correlation arises from the norm constraint. The uncertain mode shape has a deviation that is a linear combination of vectors orthogonal to the most probable mode shape.

In an overall sense the uncertainty in the mode shape as reflected by δ_Φ in (6) depends on all governing scales. Smaller ζ gives smaller δ_Φ as a result of higher dynamic amplification. Increasing N_c (normalized data length) or κ (bandwidth factor) decreases δ_Φ, since more information is used for identification. Reducing ν decreases δ_Φ, which is also intuitive. The effect of n on δ_Φ is less systematic and requires more explanation. As discussed before, S is an increasing function of n and so $\nu = S_e / S$ is a decreasing function of n. The overall effect of n on δ_Φ will be decided by the term $(n-1)\nu$, which depends on the additional dof added to the existing measured set. For example, when the mode shape value at all measured dofs are identical, then $\Phi_j = 1/\sqrt{n}$, $S \propto n$ and $\delta_\Phi^2 \propto (n-1)\nu = 1 - 1/n$, a slowly increasing function for moderate n. In this case including the additional dof will not increase the overall mode shape uncertainty significantly. On the other hand, if the additional dof has zero mode shape value, S and hence ν does not change and so $\delta_\Phi^2 \propto (n-1)\nu$ increases linearly. In general, adding dofs

with a smaller (bigger) mode shape value than the existing ones will tend to increase (decrease) the overall mode shape uncertainty.

3. Practical implications

In this section we discuss some implications of the uncertainty laws with regard to performing ambient vibration tests. We shall focus on the following issues: the governing uncertainty, the required data duration, the required modal s/n ratio and the measured dofs.

3.1. Governing uncertainty

The dependence of the posterior c.o.v. of a modal parameter on ζ has important implications on how difficult it can be estimated in practice. In particular, δ_f^2 and δ_Φ^2 are proportional to ζ, while on the contrary δ_ζ^2 is inversely proportional to ζ. For small ζ encountered in applications, say, 0.5%~5%, this means that the posterior uncertainty in the damping ratio is much larger than that in the natural frequency or mode shape, and so its accuracy requirement is likely to govern planning decisions, e.g., the required data length.

Some intuitive explanations for the dependence on ζ are in order. Small ζ implies that resonance oscillations decay slower and stay longer in the data. In the frequency domain, the resonance peak is more pronounced and sharper, giving better accuracy in the natural frequency. Correspondingly, the quality of mode shape also improves because the resonance oscillations dominate the measured vibration signal. On the other hand, a small damping means that the structure dissipates a small amount of energy and so in the diminishing limit it becomes impossible to identify the damping to the same relative accuracy, in the presence of uncertainty arising from the unknown loading that confuses energy balance. Note that the absolute uncertainty of ζ does reduce as ζ decreases, as evidence from its posterior standard deviation, being $\sqrt{\zeta / 2\pi N_c B_\zeta}$. It is just that

reduction rate is slower than its magnitude ζ, and so on a relative basis the uncertainty increases.

3.2. Data length requirement

As mentioned before, in practice the required data duration is likely to be governed by the posterior uncertainty in the damping ratio. The required data length as a multiple of the natural period to achieve a given posterior c.o.v. δ_ζ is given by

$$N_c = [2\pi\zeta B_\zeta(\kappa)\delta_\zeta^2]^{-1} \tag{11}$$

To give a rule-of-thumb, assume a damping ratio of 1% and a bandwidth factor of $\kappa = 6$ i.e., $B_\zeta \sim 60\%$. The required data length is then $N_c \approx 27.5/\delta_\zeta^2$, say,

$$N_c \approx 30/\delta_\zeta^2 \qquad (\zeta = 1\%, \kappa = 6) \tag{12}$$

This means that 300 natural periods are required to achieve a moderate posterior c.o.v. of $\delta_\zeta = 30\%$; 750 periods for $\delta_\zeta = 20\%$; and 3,000 periods for $\delta_\zeta = 10\%$. The corresponding c.o.v.s of the natural frequency are 0.67%, 0.27% and 0.067%, which are negligible. Smaller damping or bandwidth requires longer data length.

The value suggested in (12) is the minimum data length based on accuracy requirement. In practice it will need to be traded off with other practical constraints. When little is known about the existence of a mode in a frequency band one may increase (e.g., double) the data duration to get a clearer picture of the spectrum for deciding the number of modes in the band. On the other hand, there are situations that limit the data duration and hence the identification accuracy. For example, super-tall buildings (height >300m) have a natural period in excess of 5 seconds. Assuming 1% damping, it requires over 4 hours to achieve $\delta_\zeta = 10\%$. This duration is too long that significantly weakens the stationarity assumption in the stochastic modal excitation and the time invariance assumption of modal properties, giving rise to modeling errors that may invalidate the formulation. Wind loads during typhoons can change by orders of magnitude in a matter of an hour. Correspondingly, the damping ratio can change significantly over such period, due to amplitude dependence. In view of this, for super-tall buildings a c.o.v. of $\delta_\zeta = 30\%$

would be a reasonable accuracy to aim at, requiring about half an hour data [8][9]. This may put a practical limit on the precision of field evidence for wind effects on structures.

3.3. Signal-to-noise requirement

Excluding the mode shapes, the uncertainties of the modal parameters are insensitive to the modal s/n ratio, provided that it is sufficiently large. There is a limit to which improving the quality of data channel can improve the accuracy of modal parameters. The measurement noise just needs to be small enough, at which point further reducing it has little or no improvement on accuracy of the identified modal parameters.

From experience with field data, even for a small number of dofs, say, $n = 2$, it is common to have a PSD of modal force S greater than $100(\mu g)^2 / Hz$ for the first few fundamental modes under normal wind condition for civil engineering structures (e.g., buildings, bridges, floor slabs). With properly controlled data channels a prediction error PSD of $S_e \sim 100(\mu g)^2 / Hz$ (or lower) can be readily achieved. This gives a n/e ratio of $\nu \sim 1$ and a modal s/n ratio of $\gamma \sim 25$ for 1% damping. This value can be higher with more measured dofs, stronger environmental excitation, or smaller damping. The mode shape uncertainty in this case is, for $n = 2$, about $\delta_\Phi^2 \sim 1\% N_c^{-1}$. This indicates that the mode shape uncertainty can be easily reduced to an acceptable level in a cost-effective manner with a reasonable data length.

3.4. Measured dofs

The primary requirement of the measured dofs is to capture the modes of interest so that they are identifiable with desired details in the measured mode shapes. In terms of identification uncertainty of well-separated modes, the influence of the measured dofs is essentially captured in the modal s/n ratio (γ in (10)), as discussed in Section 2.4 and 2.5. When the modal s/n ratio is small, increasing the number of measured dofs may help reduce identification uncertainty by virtue of increasing the modal s/n ratio through the PSD of modal force. The rate at which the modal s/n ratio can be increased depends on

the mode shape value of the additional measured dof. Equation (9) shows that the PSD of modal force is proportional to the sum of squares of the mode shape values at the measured dofs. The latter can be taken as one simple objective for determining sensor location when the modal s/n ratio is small, trading off with other constraints. When the existing dofs can already capture the mode with a good modal s/n ratio, the posterior uncertainty in the modal parameters is insensitive to the measured dofs (number and location). In this case deploying additional sensors is not a cost-effective strategy to improve the accuracy in the natural frequency or damping ratio. Of course, quite often the number of measured dofs is increased simply to produce a detailed mode shape covering more locations of the structure.

4. Verification with field data

In this section we verify the uncertainty laws using ambient vibration data obtained from field tests. Three structures are considered. The first structure is a segment of the footbridge situated at the entrance of the City University of Hong Kong (see Figure 4(a)). Ten accelerometers were deployed for measuring the vertical acceleration at ten locations, as shown in Figure 4(b). The channel noise floor was about 1 $\mu g / \sqrt{Hz}$. The data was originally acquired at 2048Hz and then later decimated to 128Hz for analysis. It was obtained at midnight where human activity on the bridge was minimal.

Figure 5(a) shows the singular value spectrum (square-root of eigenvalues of the spectral density matrix) based on 5 minutes of ambient vibration data. The horizontal bars show the band that can normally be selected without incurring much modeling error. It should be noted that the singular value spectrum (which is smoothed by averaging) is referred here only for visualizing the spectral peaks. It is not involved in the Bayesian formulation nor modal identification calculations. Rather, the raw (complex-valued) FFT of ambient vibration data within the selected frequency band is directly involved in the Bayesian modal identification process.

Figure 5(b) shows three vertical modes identified using the FFT data within the respective bands. These give a basic idea (nominal case) of the modes that we shall focus on when verifying the uncertainty laws. Our discussion shall center around three aspects, namely the accuracy of the uncertainty laws; the variation of posterior uncertainty with data length; and the variation of posterior uncertainty with the bandwidth utilized for identification. In all cases, the value of the posterior c.o.v. according to the uncertainty laws is calculated by replacing the parameters in the formula with the MPVs (most probable value) identified from the data under question.

4.1. Effect of data length

We first examine the variation of posterior c.o.v. with data length. For this purpose we determine the most probable value (MPV) and posterior c.o.v. of the modal parameters using different data durations, being 0.5, 1, 2, 4, 8, 12, 15 minutes. In each case the bandwidth used for each mode is set as $f_0(1 \pm \kappa\zeta_0)$, where $\kappa = 6$; f_0 and ζ_0 are MPV of the natural frequency and damping in the nominal case.

Figure 6 shows the variation of the c.o.v. of modal parameters with the data length in terms of either N_c or N_f, whichever is relevant. In each plot, the markers show the 'exact' values of posterior c.o.v. computed using the fast Bayesian FFT algorithm [16], with circle, square and diamond corresponding to the first, second and third mode. The lines show the results of the uncertainty laws, with solid line, dashed line and center line corresponding to the first, second and third mode. The same notation will be used for other plots later.

Figure 6(a) shows a decreasing trend of posterior c.o.v. with the data length, as expected. Ideally, according to the uncertainty laws, for each mode the results should form a straight line with a slope of -1/2. The observed deviation from a straight line is due to the fluctuation in the MPV of modal parameters as different data duration is used. The c.o.v. of the natural frequency and damping ratio are of different order of magnitude. The c.o.v. of f increases with the mode number, while a reversed trend is observed in ζ. This is

simply because for this bridge the damping ratio decreases with the mode number. In Figure 6(b), the c.o.v. of S and S_e are plotted with N_f instead of N_c because their uncertainty laws depend directly on N_f rather than N_c. For both S and S_e the results for different modes gather together around a line. This is expected because their uncertainty laws do not depend on either f or ζ.

Figure 6(c) shows the variation of the principal variances of the posterior covariance matrix of the mode shape, $\mathbf{C}_\Phi \in R^{n\times n}$. Since $n=10$ in this example, for each mode there are 9 non-zero principle variances. According to the uncertainty law these principle variances are identical and are inversely proportional to N_c. This is reflected by the three straight lines for the three modes, which display a slope of -1 on the log-log plot. The exact values of the principal variances for each mode, on the other hand, are very close to each other, as evidence from the markers of each mode overlapping each other. The difference in the principal variances for the three modes is due to the difference in the damping ratio and the n/e ratio.

Figure 6(d) shows the variation of the overall mode shape uncertainty with N_c, in terms of the complement of expected MAC, i.e., $(1-\overline{\rho})$. For each mode, the results display approximately a straight line on the log-log plot with a slope of -1. This can be expected from (5) and (6), since $\overline{\rho} \sim 1-\delta_\Phi^2/2$ for small δ_Φ^2 and $\delta_\Phi^2 \propto N_c^{-1}$, giving $1-\overline{\rho} \propto N_c^{-1}$. Overall, Figure 6(a)-(d) indicate that the uncertainty laws give a good approximation of the posterior c.o.v.

4.2. Effect of bandwidth used

We next investigate the variation of posterior c.o.v. with the bandwidth factor κ. For this purpose we fix the data duration to be 5 minutes (same as the nominal case). We then identify the modal parameters and determine their posterior c.o.v. by utilizing the FFT data confined to the frequency band $f_0(1\pm\kappa\zeta_0)$, where κ shall now be varied from 3 to 10 at an increment of 0.5. This represents a situation where the total amount of

information (data length) is fixed and one improves knowledge regarding the modal parameters by utilizing a wider bandwidth, until practically all relevant information are exhausted. Of course, in reality one would directly use the widest possible bandwidth without increments. Note that the frequency spacing of the FFT data is $1/(5\times 60) = 3.33\times 10^{-3}$ Hz throughout.

Figure 7(a)-(f) show the variation of posterior uncertainties with κ for different modal parameters. Although Figure 7(d)-(f) display an intuitive decaying trend with κ, the same is not true for Figure 7(a)-(c), which is somewhat counter-intuitive. This is partly due to the fact that MPV of modal parameters changes as the bandwidth used increases. On the other hand, the calculated value of posterior uncertainty need not decrease with the amount of available data because the additional data may not agree with the ones that are already used. The fact that the results for S_e, principal variances and $(1-\overline{\rho})$ show a decreasing trend suggests that these properties are identified quite consistently regardless of the bandwidth used. Regardless of the trend, the uncertainty laws generally give a good approximation of the exact values.

4.3. Effect of measured dofs

To investigate the variation of posterior uncertainty with the number of measured dofs n, we fix the data duration to be 5 minutes and the bandwidth factor to be 6 (same as the nominal case). We then identify the modal properties and determine their posterior c.o.v. based on data sets with an increasing number of measured dofs. The number of dofs of these data sets ranges from 2 to 10 at an increment of 2, corresponding to different sets of measured dofs {1,2}, {1,2,3,4}, {1,2,3,4,5,6},…,{1,…,10}.

Figure 8(a)-(f) show the results for different modal parameters. It is seen from Figure 8(a)-(c) that there is no dependence of the c.o.v. of f, ζ or S on n. The decreasing trend of the c.o.v. of S_e in Figure 8(d) is also consistent with its uncertainty law. The principal variances in Figure 8(e) and $(1-\overline{\rho})$ in Figure 8(f) do depend on n in a systematic manner. The fluctuation in the overall mode shape uncertainty for small n in

Figure 8(e) is due to the fluctuation in the identified damping ratios (MPV). These results indicate that when the modal s/n ratio is sufficiently high increasing the number of measured dofs does not significantly improve the identification quality. This is consistent with the uncertainty laws.

4.4. Other structures

The uncertainty laws have also been verified with other structures. Two cases are presented here as a supplement to the CityU footbridge. The first one is a super-tall building in Hong Kong (310m high, $40\text{m} \times 40\text{m}$ in plan) with 1.5 hours of ambient data (sampling rate 128Hz) under normal wind conditions. The data was measured at 8 horizontal dofs covering four corners on the roof of the building. The second structure is the UCLA Doris and Louis Factor Health Science Building (66m high, $126\text{m} \times 73\text{m}$ in plan). The data was sampled at 100Hz and lasted for 2 hours. It corresponds to the NS horizontal dofs located on the 1/F to 16/F on the East Wall of the building. Further details of the building can be found in [18]. Figure 9 and Figure 10 show the results for the super-tall building; Figure 11 and Figure 12 for the Factor building. They are self-explanatory. Conclusions similar to the CityU footbridge can be drawn. In general, the uncertainty laws give a good approximation of the exact values.

Finally, Table 1 summarizes the scales in the examples (nominal case) for reference, which gives an idea of the typical scales that may be encountered in reality. The modal s/n ratios reported correspond to using all measured dofs in each example. They are generally quite high. When a small number of dofs is used for identification the modal s/n ratio may be smaller than that reported in the table. Note that the data were collected under normal ambient conditions where the environmental disturbance was not particularly intense. Of course, the data channels were all properly controlled to have good quality with micro-g resolution sensors. The n/e ratio ranges from 1% to 100. The bandwidth factor is roughly about 10.

5. Conclusions

These two companion papers have shown that for well-separated modes, small damping ratio and sufficient data, the posterior uncertainties of the modal parameters to the leading order take on a remarkably simple form. The uncertainty laws are summarized in (1) and (3). They are governed by the following dimensionless parameters: the damping ratio, the bandwidth factor, the noise-to-environment ratio, and the data length in terms of either the number of natural periods or the number of frequency ordinates in the selected band. The modal parameters are asymptotically uncorrelated, with the exception of the correlation between the damping ratio and the PSD of modal force.

When the modal s/n ratio is sufficiently high, further increasing it has no leading order effect on reducing the posterior uncertainties of the natural frequency or damping ratio. Increasing the number of measured dofs has no leading order effect, either. The required data length is likely to be governed by the accuracy requirement in the damping ratio. A rule of thumb has been suggested in (12), which should be traded off with practical constraints.

The uncertainty laws have been verified with field test data obtained from field test data. They generally give a good approximation. Again, we emphasis that the uncertainty laws are intended to give insights and provide guidance for planning ambient vibration tests or drafting specifications. After all, given the data the exact value of the posterior uncertainties can be calculated quickly using the fast algorithms [16][17][19].

As the Bayesian FFT approach allows full use of information in the selected frequency band of the data for given modeling assumptions, the uncertainty laws obtained in this work represent the lower limit of uncertainty that can be achieved by any method, including Bayesian and non-Bayesian methods. In the latter, uncertainty is interpreted as the ensemble variability of the estimates in a frequentist sense when there is no modeling error [5].

The reader is cautioned that the uncertainty laws describe only the leading order of the remaining uncertainty of the modal parameters for given data and modeling assumptions. They do not necessarily describe the variability of the identified values (MPV) from one data set to another. Such ensemble variability defined in a 'frequentist' manner is associated not only with identification uncertainty but also with other factors such as modeling error and variability in the structure or environment over different data sets. The (Bayesian) posterior uncertainty is for quality control rather than describing ensemble variability. The Bayesian and frequentist measure of uncertainty are related and their use are complementary [5].

Finally, answers to the questions posed in the abstract are in order, based on the context in this work. The reader should note that they can be controversial, especially when the context of application differs; the answers are as good as the assumptions.

Question 1. To estimate the damping ratio to within 30% of posterior coefficient of variation (c.o.v), what is minimum duration required?

Answer: It depends on the damping ratio ζ. Assuming a good modal s/n ratio, the required data length in terms of the number of natural periods is roughly $N_c \approx 3/\zeta$. E.g., a mode with a period of 2 seconds and 1% damping requires at least $2 \times 3/0.01 = 600$ seconds, i.e., 5 minutes.

Question 2. Will deploying an additional accelerometer significantly improve the accuracy in damping (or frequency)?

Answer: It depends on the modal s/n ratio. If it is small (say <10) then deploying an additional sensor may effectively reduce identification uncertainty by virtue of increasing the modal s/n ratio, but the extent depends on the mode shape value of the additional sensor location (higher the better). If the modal s/n ratio is already sufficiently large, the additional sensor will not significantly improve the accuracy in the damping (or frequency).

6. Acknowledgements

This paper is partly supported by General Research Fund 9041758 (CityU 110012) from the Research Grants Council of the Hong Kong Special Administrative Region, China. The data from the CityU bridge was obtained with the assistance of Dr Feng-Liang Zhang and Dr Yan-Chun Ni while they were PhD students at CityU. The data from the super-tall building was obtained through Collaborative Research Project (URN 05/81d) with Ove Arup and Partners Hong Kong Ltd (OAP). The author would like to thank Dr. Alex To, Associate at OAP, for providing wind engineering expertise and logistics support of field tests with the super-tall building. He would also like to thank Dr. Kohler for providing the UCLA Factor Building data and generous assistance during the process. Constructive comments by anonymous reviewers are gratefully acknowledged.

7. References

[1] G.C. Goodwin, R.L. Payne, Dynamic system identification: experiment design and data analysis. London: Academic Press, 1977.

[2] P.H. Kirkegaard, R. Brincker, On the optimal locations of sensors for parametric identification of linear structural systems, Mechanical Systems and Signal Processing, 8 (1994) 639–47.

[3] C. Papadimitriou, Optimal sensor placement methodology for parametric identification of structural systems, Journal of Sound and Vibration, 278 (4-5) (2004) 923-947.

[4] T.H. Yi, H.N. Li, Methodology Developments in Sensor Placement for Health Monitoring of Civil Infrastructures, Journal of Distributed Sensor Networks, 2012, Article ID 612726, doi:10.1155/2012/612726.

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

[6] A.P. Jeary, Damping in structures, Wind Engineering and Industrial Aerodynamics, 72 (1997) 345–355.

[7] N. Satake, K. Suda, T. Arakawa, A. Sasaki, A. Tamura, Damping evaluation using full-scale data of buildings in Japan, Journal of Structural Engineering, 129 (2003) 470–477.

[8] S.K. Au, F.L. Zhang, P. To, Field observations on modal properties of two tall buildings in Hong Kong, Wind Engineering and Industrial Aerodynamics, 101 (2012) 12-23.

[9] 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.

[10] H. Koruk, K. Y. Sanliturk, Damping uncertainty due to noise and exponential windowing, Journal of Sound and Vibration, 330(23) (2011) 5690–5706.

[11] W. Fladung, R. Rost, Application and correction of the exponential window for frequency response functions. Mechanical Systems and Signal Processing 11(1) (1997) 23–36.

[12] H. Jokinen, J. Ollila, O. Aumala, On windowing effects in estimating averaged periodograms of noisy signals, Measurement, 28 (2000) 197–207.

[13] P. J. Torvik, On estimating system damping from frequency response bandwidths. Journal of Sound and Vibration,330 (2011) 6088–6097

[14] S. K. Au, F. L. Zhang and Y. C. Ni, Bayesian operational modal analysis: theory, computation, practice, Computers and Structures, (2013). In print. http://dx.doi.org/10.1016/j.compstruc.2012.12.015.

[15] S.K. Au, F.L. Zhang, On assessing posterior mode shape uncertainty in ambient modal identification, Probabilistic Engineering Mechanics, 26 (3) (2011) 427-434.

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

[17] F.L. Zhang, S.K. Au, Erratum to 'Fast Bayesian FFT method for ambient modal identification with separated modes' by Siu-Kui Au, Journal of Engineering Mechanics, ASCE, 139(4) (2013) 545-545.

[18] M.D. Kohler, P.M. Davis, E. Safak, Earthquake and ambient vibration Monitoring of the steel-frame UCLA Factor Building, Earthquake Spectra, 21(3) (2005) 715–736.

[19] 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.

List of Tables

List of Figures

Table 1. Scales in different field examples (nominal case)

Case	Mode	Modal s/n ratio γ	n/e ratio ν	Bandwidth factor κ
CityU footbridge	1	13134	0.14	10
	2	5994	1.0	19
	3	2081	106	7
Super-tall building	1	1511	37	22
	2	1376	10	14
	3	253	25	11
Factor building	1	94515	0.0044	12
	2	17350	0.12	29
	3	601	0.26	4

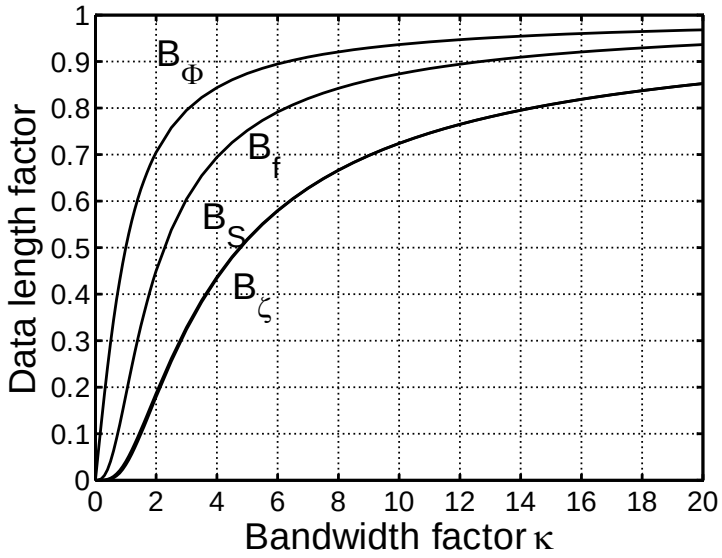


Figure 1. Data length factors

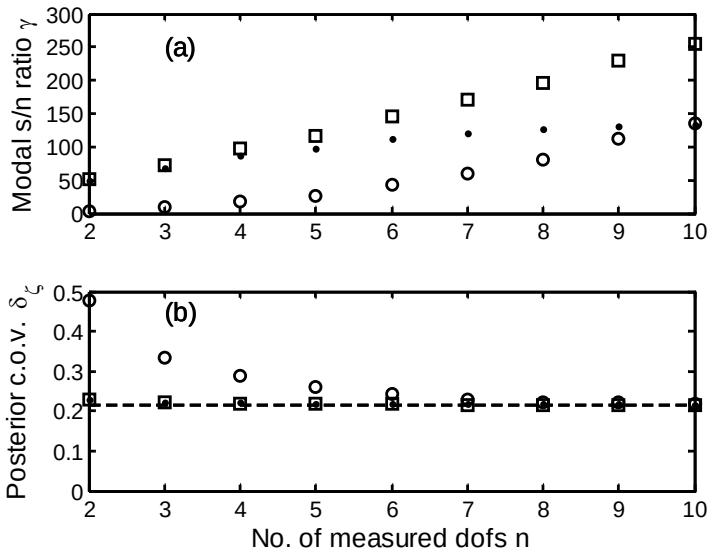


Figure 2. Modal s/n ratio (a) and posterior c.o.v. of damping ratio (b) versus the measured number of dofs, ten-storied shear building example (synthetic data). Dots – sensor from roof to 1/F; circles – sensor from 1/F to roof; squares – all sensors at roof. Dashed line in (b) – uncertainty law

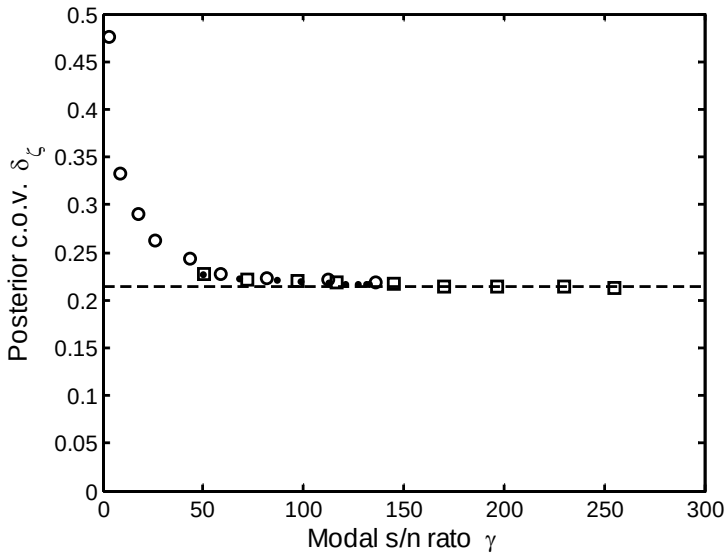
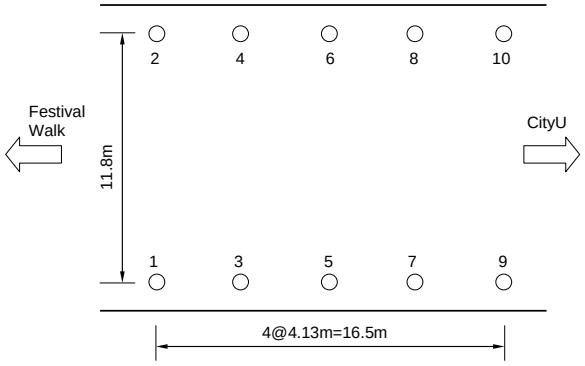


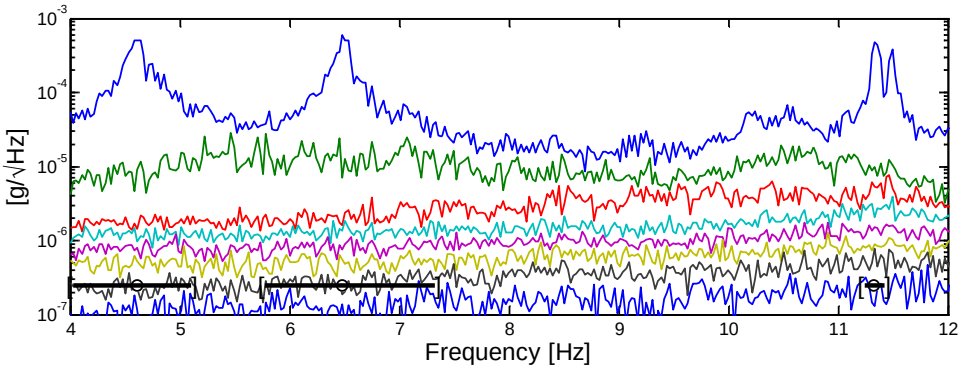
Figure 3. Posterior c.o.v. of damping versus modal s/n ratio, ten-storied shear building example (synthetic data). Dots – sensor from roof to 1/F; circles – sensor from 1/F to roof; squares – all sensors at roof. Dashed line – uncertainty law

(a) View from below

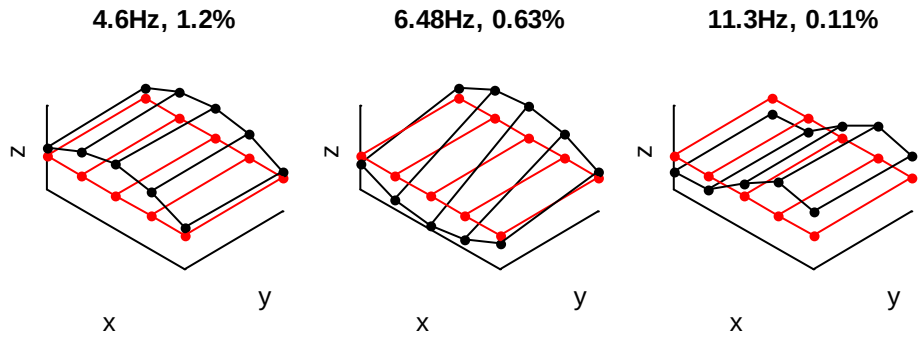


(b) Setup plan

Figure 4. CityU footbridge



(a) (Root) Singular value spectrum (averaged for viewing modes only)



(b) Identified modal properties (MPV)

Figure 5. Analysis of CityU footbridge (nominal case, 5 min. data)

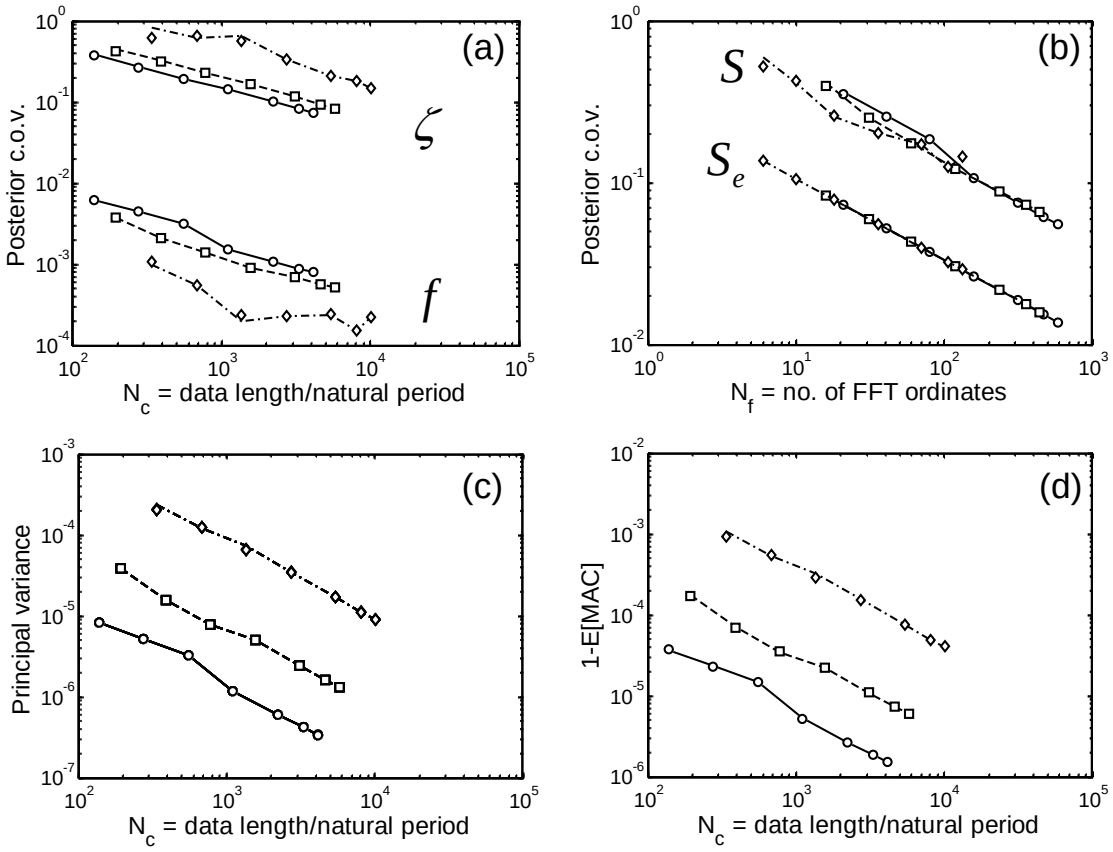


Figure 6. Posterior c.o.v. versus data length, CityU footbridge. Marker – exact; line – uncertainty law; circle, solid line – Mode 1; square, dashed line – Mode 2; diamond, center line – Mode 3.

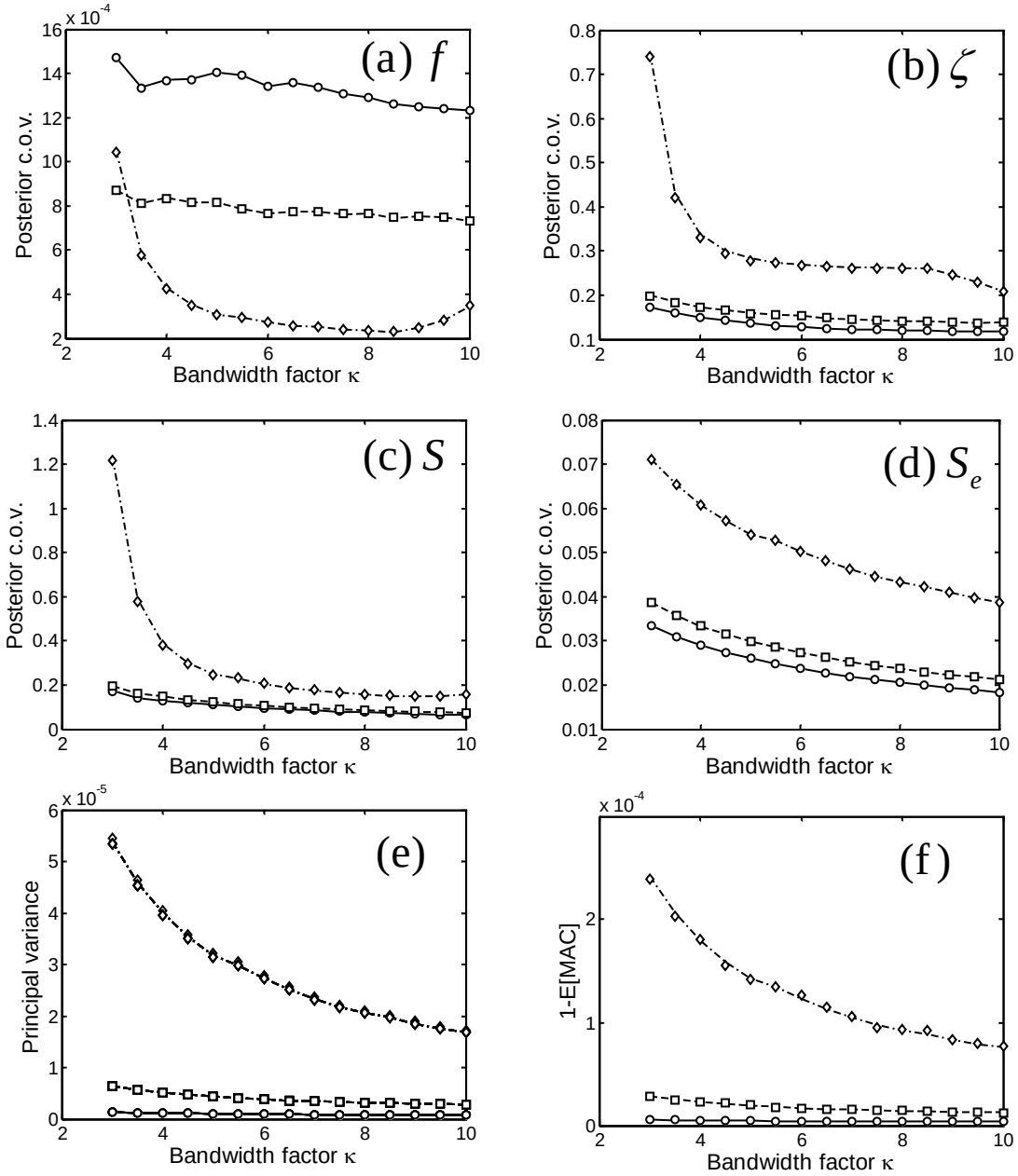


Figure 7. Posterior c.o.v. versus bandwidth factor, CityU footbridge. Legend same as Figure 6.

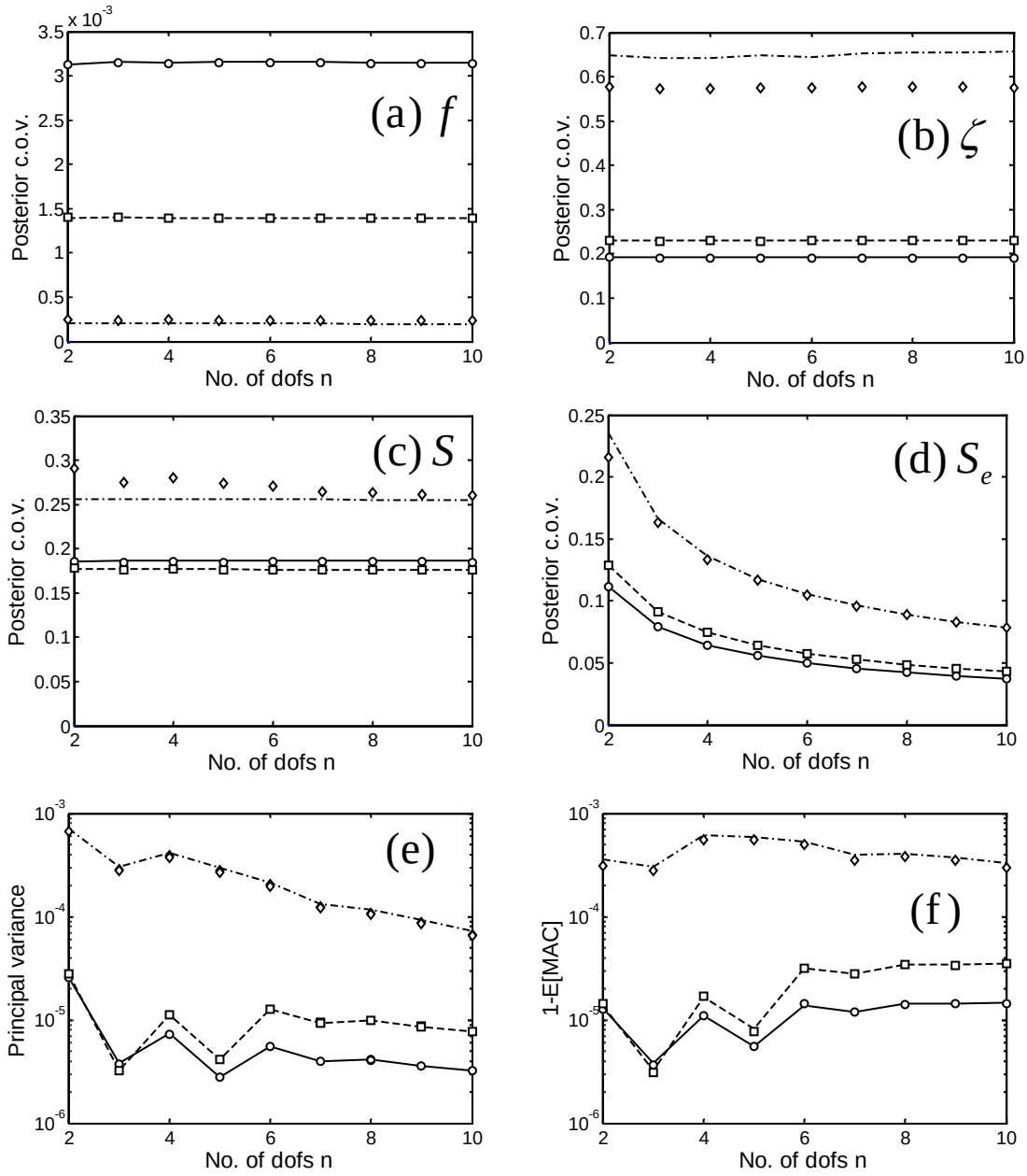
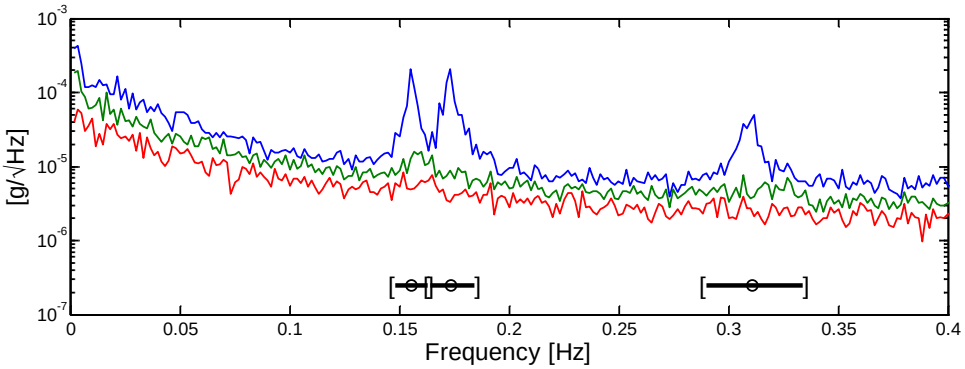
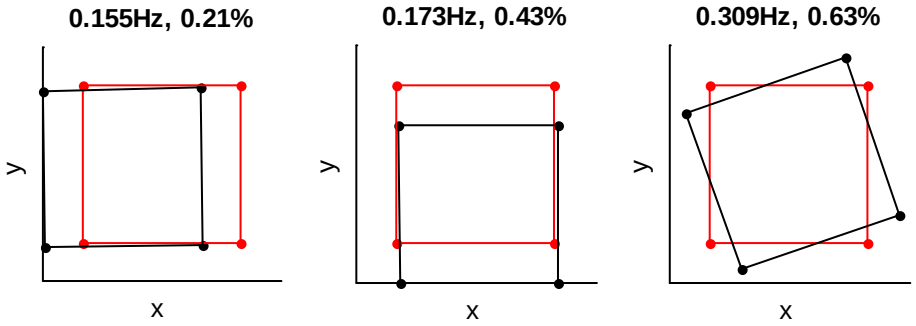


Figure 8. Posterior c.o.v. versus no. of dofs, CityU footbridge. Legend same as Figure 6.



(a) (Root) Singular value spectrum (averaged for viewing modes only)



(a) Identified modal properties (MPV)

Figure 9. Analysis of super-tall building (nominal case, 30 min. data)

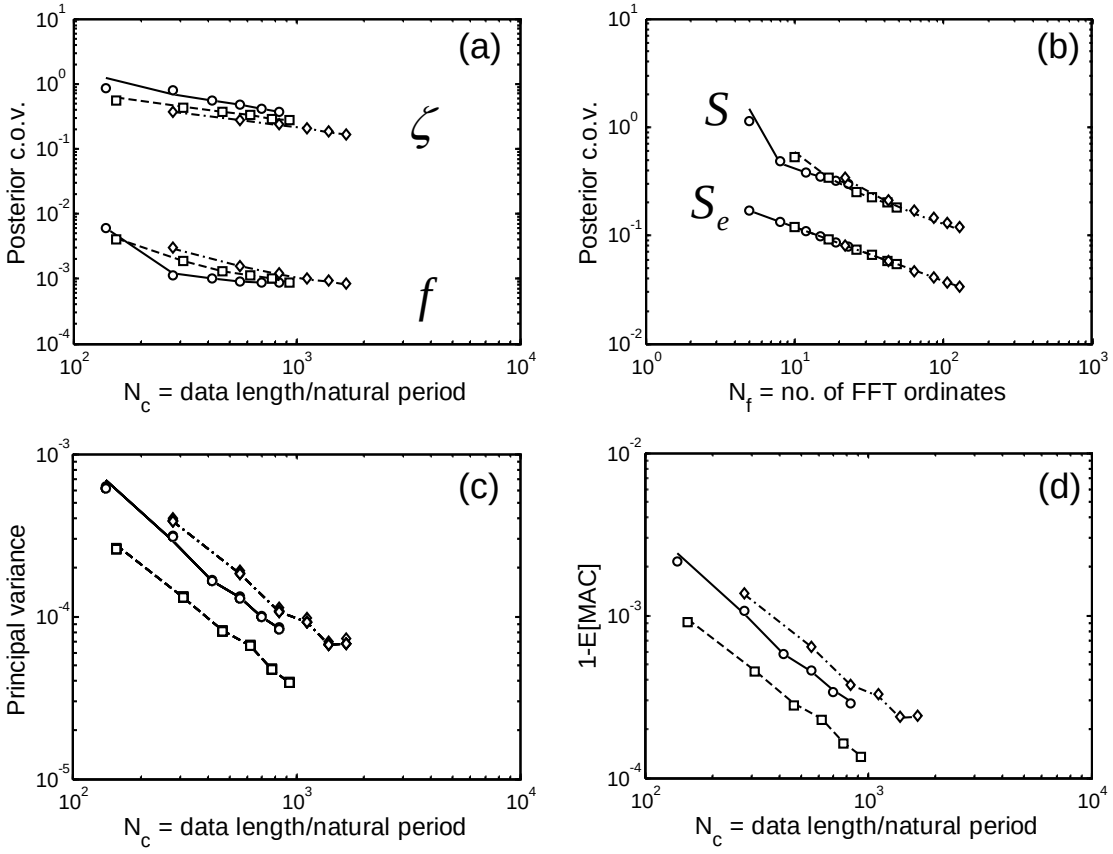
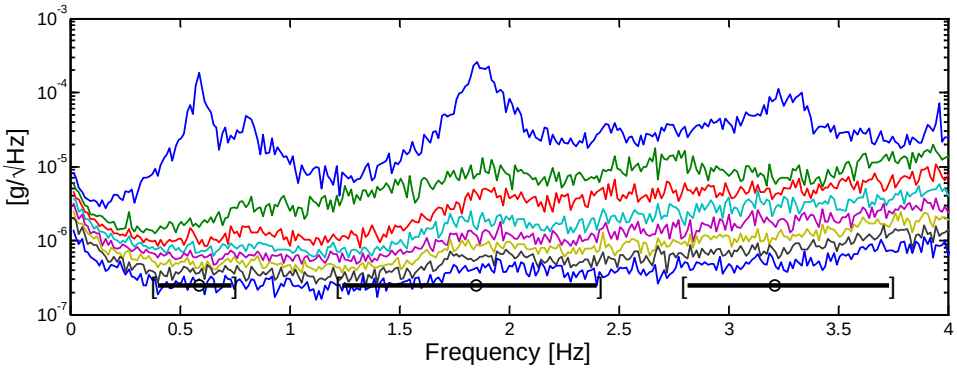
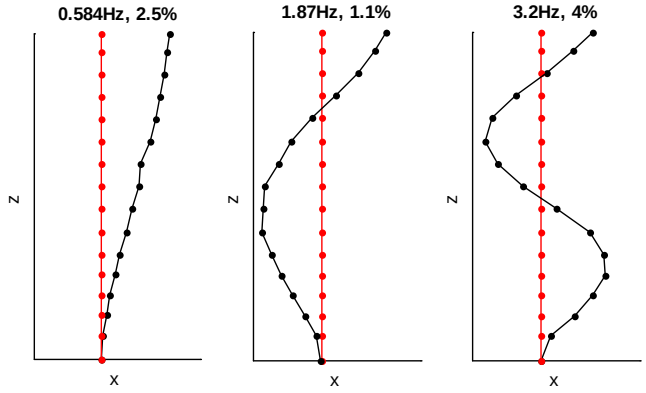


Figure 10. Posterior c.o.v. versus data length, super-tall building. Legend same as Figure 6.



(a) (Root) Singular value spectrum (averaged for viewing modes only)



(b) Identified modal properties (MPV)

Figure 11. Analysis of Factor Building (nominal case, 10 min. data)

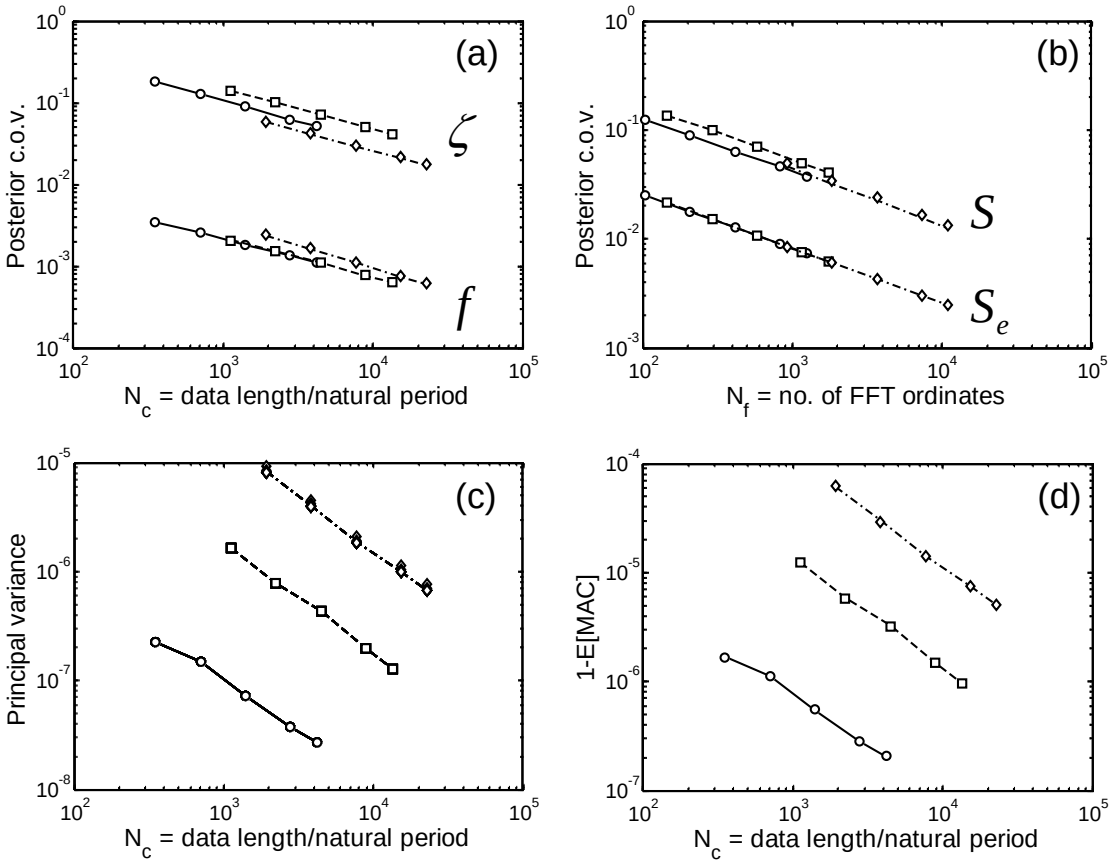


Figure 12. Posterior c.o.v. versus data length, Factor Building. Legend same as Figure 6.