Journal of Mechanical and Electrical Intelligent System (JMEIS)

Variational Bayes estimation of unknown parameter distribution in 1-DoF vibratory system which subjected to random excitation

Soichiro Takata[1, a,*] and Hirofumi Inoue[2, b]

[1]Department of Mechanical Engineering, National Institute of Technology, Tokyo College, 1220-2 Kunugida-machi, Hachioji-shi, 193-0997, Japan

[2]NEC Corporation, 1753 Shimonumabe, Nakahara-ku, Kawasaki-shi, 211-8666, Japan

*Corresponding author

[a]<takata@tokyo-ct.ac.jp>, [b]<h-inoue0822@nec.com>

Keywords: random vibration, system identification, Fokker-Planck equation, Maxwell-Boltzmann distribution, Variational Bayes inference

Abstract. This paper presents an expansion method for parameter distribution estimation based on variational Bayes inference in a linear single-degree-of-freedom system using Gaussian random vibration responses. The likelihood function of the proposed method is defined by the analytical solution of the Fokker–Planck equation derived in a previous study. The unknown parameters are estimated using the variational Bayes formula. Furthermore, numerical identifications are conducted using random responses from the results of the 4th-order Runge–Kutta method. The estimated performance of the proposed method was verified in terms of the dependence on the sample size. Benchmark tests were conducted to compare the accuracy of the variational Bayes and maximum likelihood estimations. The variational Bayes estimation exhibited higher accuracy than the maximum likelihood estimation for small sample sizes. Furthermore, a high-accuracy implementation trial was conducted with a focus on the dependence of the calculation sequence on the expected value of the variance.

1. Introduction

System identification based on random vibration responses has traditionally been used in the field of mechanical and structural vibrations, wherein autoregressive time series analysis is the most extensively employed method [1,2,3,4]. Since this method only requires output data and obtaining the input data is generally difficult in field tests, it is often used for actual field data. However, owing to its dependence on the output data, its accuracy in predicting the error variance is low. In recent years, the ARX model which uses input–output data, has been applied to solve the above problem [5,6].

Operational modal analysis (OMA) has recently been developed in the field of stochastic signal processing technologies [7,8,9,10]. Mode shape visualization and estimation of modal parameters are conducted using a frequency response based on the modal expansion, which uses white noise excitation as the input. Subsequently, the modal shape in the actual structure can be understood using OMA. This technique is available for the analysis of actual field sensor data. However, as it is based on the frequency method, it cannot be applied to large-damping structural and material systems.

In previous reports, we proposed an identification method based on maximum likelihood estimation using a probability density function [11]. This function obeys system parameters such as the spring constant, damping constant, and diffusion coefficient of the input white noise. These characteristics are unrelated to the frequency spectrum. Thus, a method based on probability density function can be applied to large damping-vibratory systems.

Furthermore, the above identification method based on maximum likelihood estimation has been applied to the hardening-spring Duffing system [12]. The formula for estimating the analytical form was not obtained because the optimization function was nonlinear with respect to the unknown parameters. Hence, a numerical solver was proposed for the optimization function, which focused on the ratio of the orders of moments.

The method based on maximum likelihood estimation uses point estimation and cannot obtain the unknown parameter distribution. If the parameter distribution is obtained, various statistics for the unknown parameters can be calculated. A framework using variational Bayes inference can obtain the unknown parameter distribution [13,14,15,16,17,18,19].

In this study, we investigated an identification method based on the variational Bayes inference in a single-degree-of-freedom (1-DoF) vibratory system that was subjected to white random excitation. First, the identification algorithm was derived based on variational Bayes inference. The proposed likelihood function was formulated using the Maxwell-Boltzmann distribution. A numerical simulation was performed using the derived identification algorithm. Moreover, benchmark tests were conducted to compare the variational Bayes and the maximum likelihood estimations. A high-accuracy implementation trial was conducted with a focus on the dependence of the calculation sequence on the expectation value of the variance.

2. Derivation of identification algorithm

2.1 Likelihood function based on Maxwell-Boltzmann distribution

The Fokker–Planck equation for unit mass in a 1-DoF system subjected to a white noise excitation force is as follows [11]:

$$\frac{\partial f(x_1,x_2,t)}{\partial t} = -x_2\frac{\partial f(x_1,x_2,t)}{\partial x_1} + (kx_1 + cx_2)\frac{\partial f(x_1,x_2,t)}{\partial x_2} + cf(x_1,x_2,t) + D\frac{\partial^2 f(x_1,x_2,t)}{\partial x_2^2}, \quad (1)$$

where $f(x_1,x_2,t)$ and D represent the probability density function of the stochastic response and the diffusion coefficient, respectively, and k and c represent the spring constant and damping coefficient, respectively. Furthermore, the displacement $x_1(t) = x(t)$ and velocity $x_2(t) = dx_1(t)/dt$. The analytical solution for the stationary Fokker–Planck equation is as follows:

$$f_s(x_1,x_2) = \frac{c\sqrt{k}}{2\pi D}\exp\left[-\frac{c}{2D}(kx_1^2 + x_2^2)\right]. \quad (2)$$

Eq. (2) contains the parameters k and D/c and the stationary distribution is represented by $f_s(x_1,x_2)$. The analytical solution for the Fokker–Planck equation is equal to the Maxwell–Boltzmann distribution [20]. Therefore, we refer to the likelihood function obtained using the analytical solution of the Fokker–Planck equation as the "Maxwell–Boltzmann likelihood" in a broad sense.

In this study, we propose an estimation method for the unknown parameters k and D/c based on variational Bayes inference [13]. The Maxwell–Boltzmann likelihood is defined by the analytical solution of the Fokker–Planck equation defined in Eq. (2) and is observed as time series data. The observed time series data contain the displacement x_1 and velocity x_2, and the datasets are $\mathbf{D} = \{(x_{1,1},x_{2,1}),(x_{1,1},x_{2,1}),\ldots,(x_{1,\alpha},x_{2,\alpha})\ldots,(x_{1,N},x_{2,N})\}$. Therefore, the likelihood is defined as follows:

$$L = \left(\frac{\tau_1}{2\pi}\right)^{N/2}\cdot\exp\left[-\sum_{\alpha=1}^{N}\frac{\tau_1 x_{1\alpha}^2}{2}\right]\cdot\left(\frac{\tau_2}{2\pi}\right)^{\frac{N}{2}}\cdot\exp\left[-\sum_{\alpha=1}^{N}\frac{\tau_2 x_{2\alpha}^2}{2}\right], \quad (3)$$

where, $\tau_1 = k\cdot c/D$ and $\tau_2 = c/D$.

2.2 Identification algorithm based on variational Bayes inference

Variational Bayes inference is based on Bayes theorem. The assumption of a conjugate distribution between the prior and posterior distributions is required for Bayes theorem. The Maxwell–Boltzmann likelihood is also incorporated into the Bayes theorem (i.e., posterior distribution = likelihood function × prior distribution).

To apply the variational Bayes formula, the mean field approximation of the posterior distribution $q(\tau_1, \tau_2)$ is calculated using the factorized posterior distributions, $q_{\tau_1}(\tau_1)$ and $q_{\tau_2}(\tau_2)$ as follows:

$$q(\tau_1, \tau_2) = q_{\tau_1}(\tau_1) \cdot q_{\tau_2}(\tau_2), \tag{4}$$

where the factorized conjugate prior distribution assumes the gamma distribution because the Maxwell–Boltzmann likelihood is composed of a Gaussian distribution. The factorized conjugate prior distributions are defined as follows:

$$p(\tau_1) = \mathrm{Gam}(\tau_1 | a_1, b_1), \quad p(\tau_2) = \mathrm{Gam}(\tau_2 | a_2, b_2), \tag{5}$$

where a_1, b_1, a_2, b_2 represent the gamma distribution parameters. According to the variational Bayes formula [13], the following relationship exists between the factorized posterior distribution and the posterior distribution based on Baye's theorem:

$$\mathrm{In}\ q_i^*(Z_i) = \mathrm{E}_{i \neq j}[\mathrm{In}\, p(\mathbf{D}, \mathbf{Z})] + \mathrm{const}, \tag{6}$$

where Z_i represent τ_1 and τ_2, respectively, $p(\mathbf{D}, \mathbf{Z})$ represents the posterior distribution based on the Bayes theorem, $\mathbf{D}$ represents the observed data, and $\mathbf{Z} = \{\tau_1, \tau_2\}$. Furthermore, $p(\mathbf{D}|\tau_1, \tau_2)$ represents the likelihood function (i.e., Maxwell-Boltzmann likelihood). By combining Eqs. (3) to (5) into Eq. (6), the factorized posterior distribution can be expressd as follows:

$$\begin{aligned}
\ln q_{\tau_1}^*(\tau_1) &= \mathrm{E}_{\tau_2}[\ln p(\mathbf{D}, \tau_1, \tau_2)] + \mathrm{const} \\
&= \mathrm{E}_{\tau_2}[\ln\{p(\mathbf{D}|\tau_1, \tau_2) p(\tau_1) p(\tau_2)\}] + \mathrm{const} \\
&= \mathrm{E}_{\tau_2}[\ln p(\mathbf{D}|\tau_1, \tau_2) + \ln p(\tau_1) + \ln p(\tau_2)] + \mathrm{const} \\
&= \mathrm{E}_{\tau_2}[\ln p(\mathbf{D}|\tau_1, \tau_2) + \ln p(\tau_1)] + \mathrm{const} \\
&= \mathrm{E}_{\tau_2}\Bigg[\ln\left\{\left(\frac{\tau_1}{2\pi}\right)^{N/2} \exp\left[-\sum_{\alpha=1}^{N} \frac{\tau_1 x_{1\alpha}^2}{2}\right] \left(\frac{\tau_2}{2\pi}\right)^{N/2} \exp\left[-\sum_{\alpha=1}^{N} \frac{\tau_2 x_{2\alpha}^2}{2}\right]\right\} \\
&\qquad + \ln \mathrm{Gam}(\tau_1 | a_1, b_1)\Bigg] + \mathrm{const} \\
&= \mathrm{E}_{\tau_2}\left[\ln\left(\frac{\tau_1}{2\pi}\right)^{N/2} \exp\left[-\sum_{\alpha=1}^{N} \frac{\tau_1 x_{1\alpha}^2}{2}\right] + \ln\left(\frac{1}{\Gamma(a_1)} b_1^{a_1} \tau_1^{a_1 - 1} \exp[-b_1 \tau_1]\right)\right] + \mathrm{const} \\
&= \mathrm{E}_{\tau_2}\left[\ln(\tau_1)^{N/2} \exp\left[-\sum_{\alpha=1}^{N} \frac{\tau_1 x_{1\alpha}^2}{2}\right] + \ln\left(\tau_1^{a_1 - 1} \exp[-b_1 \tau_1]\right)\right] + \mathrm{const} \\
&= \mathrm{E}_{\tau_2}\left[\frac{N}{2} \ln \tau_1 - \tau_1 \sum_{\alpha=1}^{N} \frac{x_{1\alpha}^2}{2} + (a_1 - 1) \ln \tau_1 - b_1 \tau_1\right] + \mathrm{const} \\
&= \mathrm{E}_{\tau_2}\left[\left(\frac{N}{2} + a_1 - 1\right) \ln \tau_1 - \left(\sum_{\alpha=1}^{N} \frac{x_{1\alpha}^2}{2} + b_1\right) \tau_1\right] + \mathrm{const.}
\end{aligned}$$

In the above equations, the factorized posterior distribution is equal to the gamma distribution $\mathrm{Gam}(\tau_1|a_{1N}, b_{1N})$ and the parameter of the factorized posterior distribution is obtained as follows:

$$a_{1N} = a_1 + \frac{N}{2}, b_{1N} = \sum_{\alpha=1}^{N} \frac{x_{1\alpha}^2}{2} + b_1. \tag{7}$$

In the same manner, the optimization of τ_2 is expressed as follows:

$$\begin{aligned}
\ln q_{\tau_2}^*(\tau_2) &= \mathrm{E}_{\tau_1}[\ln p(\mathbf{D}, \tau_1, \tau_2)] + \mathrm{const} \\
&= \mathrm{E}_{\tau_1}[\ln p(\mathbf{D}|\tau_1, \tau_2) + \ln p(\tau_2)] + \mathrm{const} \\
&= \mathrm{E}_{\tau_1}\left[\ln\left(\frac{\tau_2}{2\pi}\right)^{N/2} \exp\left[-\sum_{\alpha=1}^{N} \frac{\tau_2 x_{2\alpha}^2}{2}\right] + \ln\left(\frac{1}{\Gamma(a_2)} b_2^{a_2} \tau_2^{a_2-1} \exp[-b_2\tau_2]\right)\right] + \mathrm{const} \\
&= \mathrm{E}_{\tau_1}\left[\ln(\tau_2)^{N/2} \exp\left[-\sum_{\alpha=1}^{N} \frac{\tau_2 x_{2\alpha}^2}{2}\right] + \ln\left(\tau_2^{a_2-1} \exp[-b_2\tau_2]\right)\right] + \mathrm{const} \\
&= \mathrm{E}_{\tau_1}\left[\left(\frac{N}{2} + a_2 - 1\right)\ln\tau_2 - \left(\sum_{\alpha=1}^{N} \frac{x_{2\alpha}^2}{2} + b_2\right)\tau_2\right] + \mathrm{const},
\end{aligned}$$

where, the factorized posterior distribution is equal to the gamma distribution $\mathrm{Gam}(\tau_2|a_{2N}, b_{2N})$, and the parameter of the factorized posterior distribution is obtained as follows:

$$a_{2N} = a_2 + \frac{N}{2}, b_{2N} = \sum_{\alpha=1}^{N} \frac{x_{2\alpha}^2}{2} + b_2. \tag{8}$$

Furthermore, the unknown parameter distributions (i.e., probability density functions for k and D/c) are obtained by transforming the probability variables using the parameters of Eqs. (7) and (8), respectively.

The probability density function of the spring constant (k) is obtained as follows:

$$p(k) = \frac{\Gamma(a_{1N} + a_{2N})}{\Gamma(a_{1N})\Gamma(a_{2N})} b_{1N}^{a_{1N}} b_{2N}^{a_{2N}} \frac{k^{a_{1N}-1}}{(b_{1N} \cdot k + b_{2N})^{a_{1N}+a_{2N}}}. \tag{9}$$

The ratio between the diffusion coefficient (D) and damping constant (c) is obtained as follows:

$$p(D/c) = \frac{1}{\Gamma(a_{2N})} b_{2N}^{a_{2N}} \frac{1}{(D/c)^2} \left(\frac{1}{D/c}\right)^{a_{2N}-1} \exp\left[-\frac{b_{2N}}{(D/c)}\right], \tag{10}$$

where, $\Gamma(\cdot)$ represents the gamma function.

3. Numerical simulation

3.1 Conditions

Flowcharts for the numerical identification algorithm using variational Bayes estimation and maximum likelihood estimation are presented in Fig. 1 (a) and (b), respectively.

In both cases, a time series was generated using the 4-th order Runge–Kutta method to estimate the unknown system parameters. The spring constant was set to $k = 1.0$, the damping constant was set to $c = 0.01, 0.1, 1.0$, the variance of the input white noise excitation was set to $\sigma_w^2 = 1$, the diffusion coefficient is $D = 0.05$, the initial conditions were $x(0) = 0$ and $v(0) = 0$, and the sampling period was $\Delta t = 0.1$. Random responses were generated for $M = 10000$ samples. The identification was performed using N samples between $M - N$ and M to eliminate the stochastic transience. Here, N was varied from 10 to 300 in increments of 10.

The variational Bayes and maximum likelihood estimations were performed using the derived formula. The process was repeated for a set number of iterations and the estimated parameter values at each iteration were saved in memory, which were then averaged to obtain the final estimated parameter values.

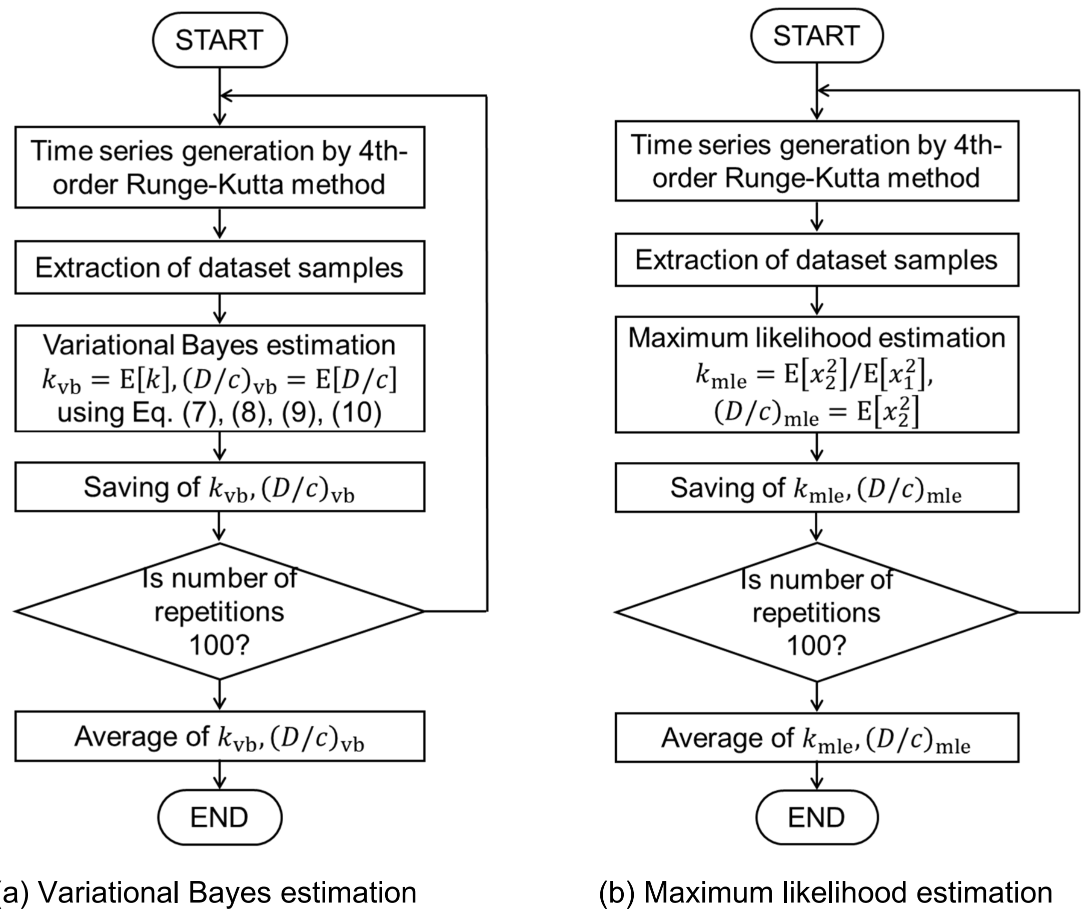
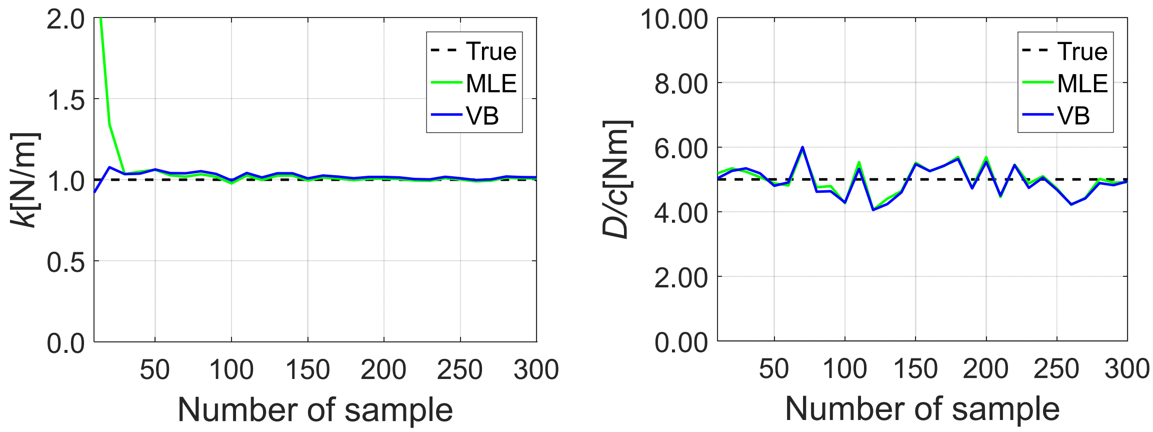


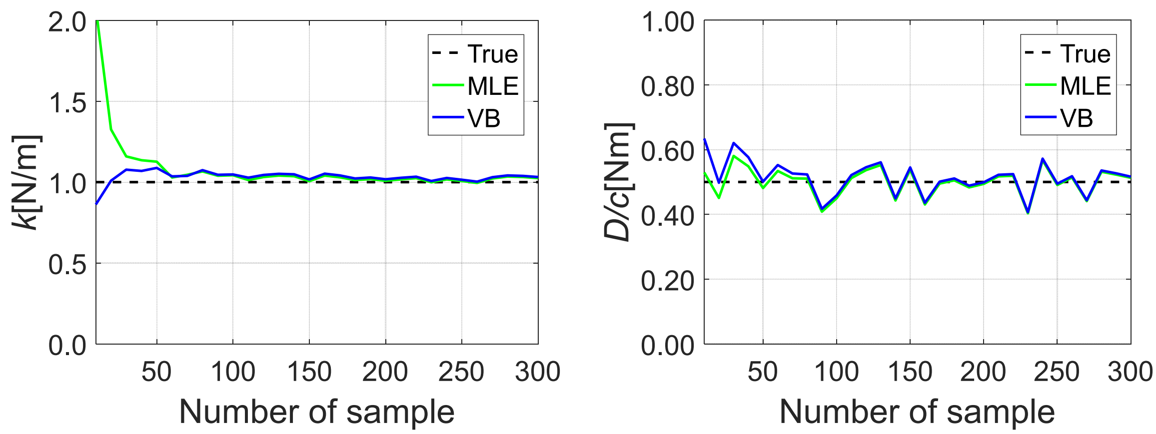
Fig. 1. Flowcharts for the numerical identification algorithm using (a) Variational Bayes estimation and (b) Maximum likelihood estimation.

3.2 Results and discussion

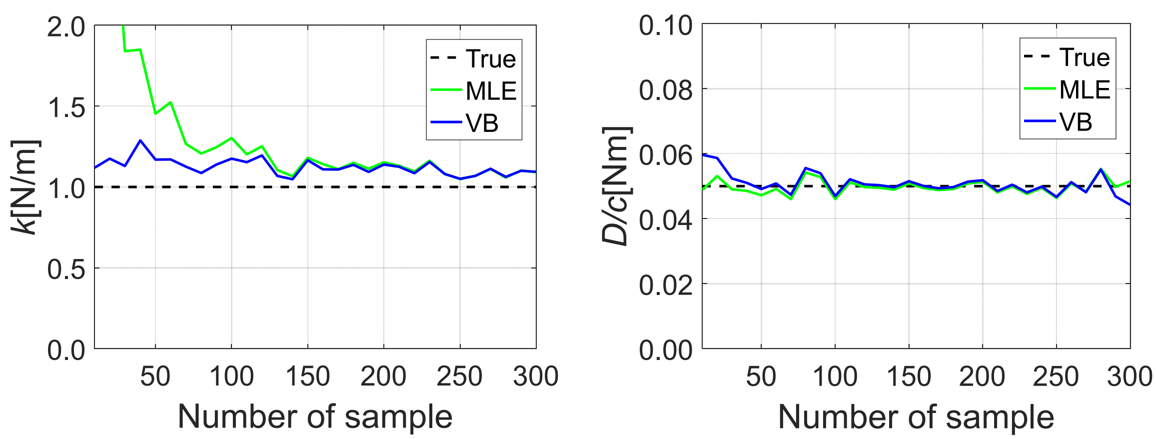
The dependencies of the estimated parameter values on the number of samples are shown in Fig. 2. The black dotted line indicates the true value, the green solid line indicates the maximum likelihood estimation (MLE), and the blue solid line indicates the case of variational Bayes (VB) estimation.



(a) $c = 0.01$ (small damping case; loss coefficient $\eta = 0.01$)



(b) $c = 0.1$ (middle damping case; loss coefficient $\eta = 0.1$)



(c) $c = 1.0$ (large damping case; loss coefficient $\eta = 1.0$)

Fig. 2. Benchmark test results in terms of sample number dependence. The left column shows the results for the spring constant (k) and the right column shows the results for the ratio diffusion coefficient (D) and the damping constant (c). The damping constant (c) is set to (a) $c = 0.01$, (b) $c = 0.1$, (c) $c = 1.0$.

The estimation values in all the cases converged with the true values as the number of samples increased. For the spring constant, the difference between the true values and the values estimated using the maximum likelihood estimation were high when number of samples were small; however, these values converged with the true values when the number of samples were large.

In contrast, the variational Bayes estimation method performed well at all sample sizes and damping constants. For the ratio between the diffusion coefficient and damping constant, no clear difference in the estimation accuracy between the variational Bayes and maximum likelihood estimations was observed.

4. High-accuracy implementation trial

A high-accuracy implementation trial was carried out. The expected value of the variance is dependent on the calculation sequence. Thus, the calculation sequence must be carefully considered for high-accuracy identification. The revised flowcharts for the numerical identification algorithm using variational Bayes estimation and maximum likelihood estimation are shown in Fig. 2 (a) and (b), respectively. The following areas were modified. The average variance was calculated before the estimation of the parameters k and D/c. Consequently, the dependance of the estimated parameter values on the sample size decreased. Therefore, in the high-accuracy implementation trial, $\mathrm{E}[x_1^2]$ and $\mathrm{E}[x_2^2]$, calculated using Eqs. (7) and (8), were averaged before the estimation of the parameters k and D/c.

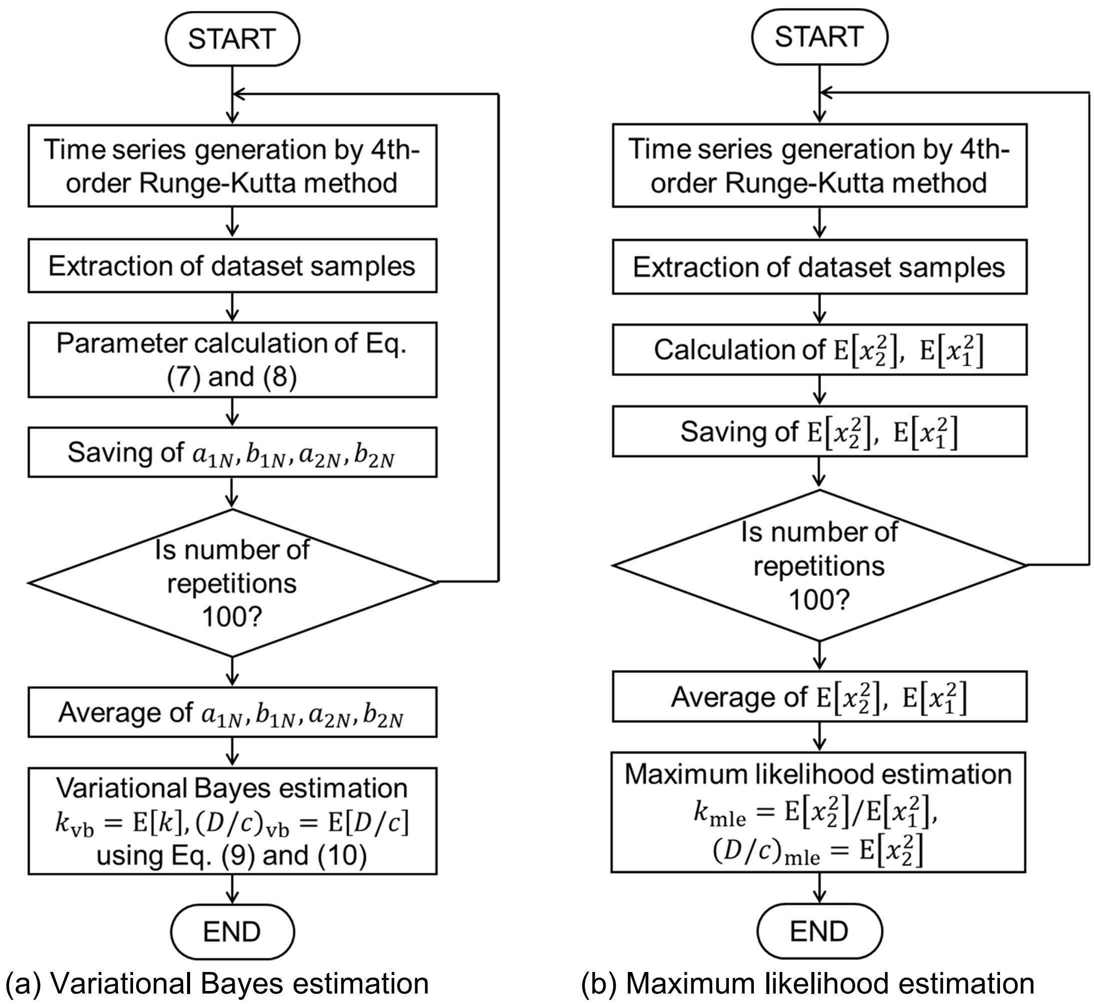
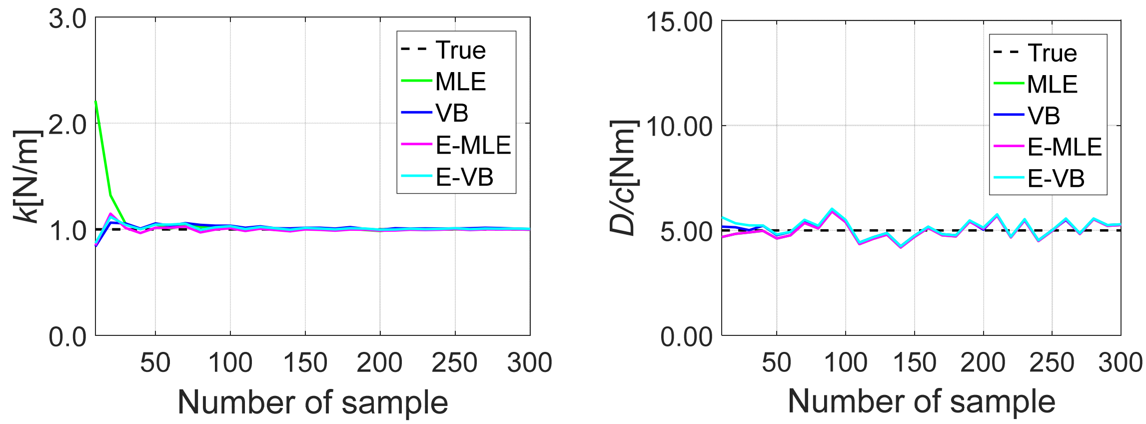
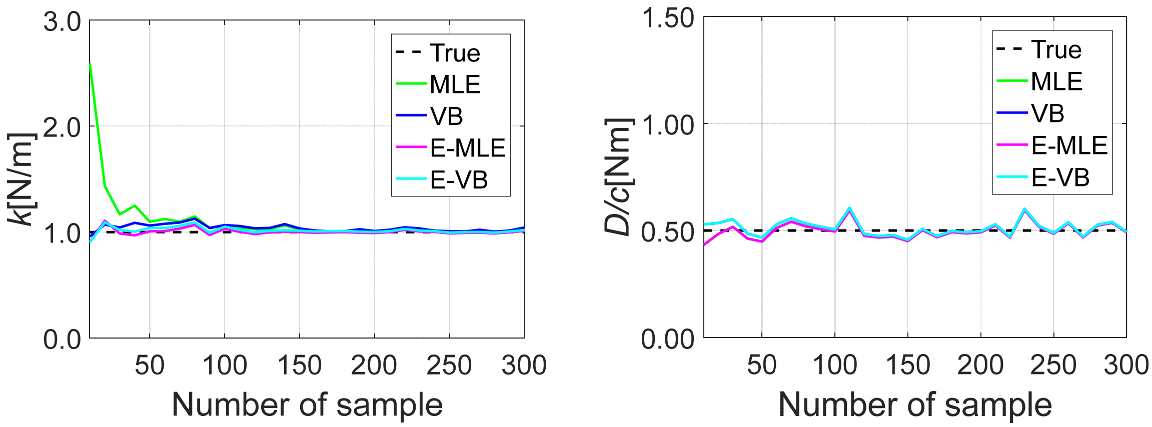


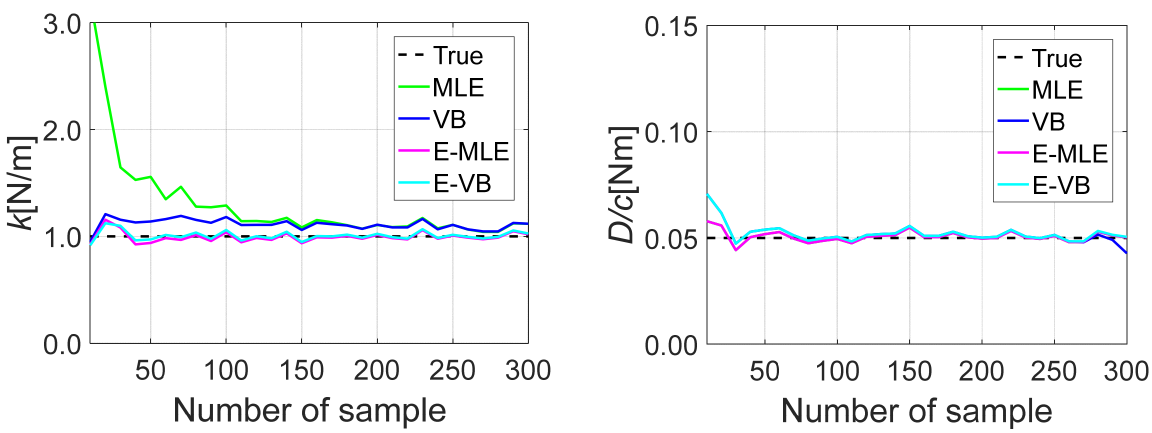
Fig. 3. Flowcharts for the numerical identification algorithm using (a) Variational Bayes estimation and (b) Maximum likelihood estimation for high-accuracy implementation.



(a) $c = 0.01$ (small damping case; loss coefficient $\eta = 0.01$)



(b) $c = 0.1$ (middle damping case; loss coefficient $\eta = 0.1$)



(c) $c = 1.0$ (large damping case; loss coefficient $\eta = 1.0$)

Fig. 4. Benchmark test results for the high-accuracy implementation trial. The left column shows the results for the spring constant (k) and the right column shows the results for the ratio diffusion coefficient (D) and the damping constant (c). The damping constant (c) is set to (a) $c = 0.01$, (b) $c = 0.1$, and (c) $c = 1.0$.

The dependencies of the estimated parameter values on the number of samples for the high-accuracy implementation trial are shown in Fig. 4. The black dotted line indicates the true value, the green solid line indicates the maximum likelihood estimation (MLE), the blue solid line indicates the variational Bayes (VB), the magenta solid line represents the revised MLE (effective-MLE; E-MLE), and the light blue solid line indicates the revised VB (effective-VB; E-VB).

The estimated parameter values converged with the true values in all the cases the number of samples increased. For the spring constant, the estimation accuracy of E-MLE and E-VB were higher than those of MLE and VB. Furthermore, the values estimated using E-VB converged with those estimated using E-MLE. Thus, the effectiveness of the proposed high-accuracy implementation was verified for the estimation of the spring constant parameter. However, in the case of the ratio between the diffusion coefficient and damping constant, the values estimated using MLE and VB were in agreement with those estimated using E-MLE and E-VB. Thus, the effectiveness of the proposed high-accuracy implementation was not observed in the case of estimating the ratio the between diffusion coefficient and damping constant.

5. Conclusions

In this paper, we have presented identification method based on variational Bayes estimation using the Maxwell–Boltzmann likelihood. The following results were obtained.

(1) The identification algorithm based on variational Bayes inference was formulated using the Maxwell–Boltzmann likelihood.
(2) Benchmark tests on the estimation accuracy were conducted for the maximum likelihood estimation. The variational Bayes estimation exhibited higher accuracy than the maximum likelihood estimation when the sample size was small.
(3) A high-accuracy implementation trial was conducted by considering the dependence of the calculation sequence on the expected value of the variance. The effectiveness of the proposed implementation was verified in the case of the spring constant estimation.

References

[1] K. Suzuki and A. Nakajima, “Practical method to estimate damping ratio based on the characteristics of AR spectral analysis technique”, ***Transactions of the Japan Society of Mechanical Engineers, Series C***, Vol. 48, No. 433, pp. 1389–1397, 1982 (in Japanese).

[2] K. Suzuki and Y. Harada, “Analysis of structural nonstationary behavior and dynamic system identification based on the auto-regressive model fitting”, ***Transactions of the Japan Society of Mechanical Engineers, Series C***, Vol. 50, No. 458, pp. 2013–2020, 1984 (in Japanese).

[3] H. Sohn and C. R. Farrar, “Damage diagnosis using time series analysis of vibration signals”, ***SMART MATERIALS AND STRUCTURES***, Vol. 10, pp. 446–451, 2001.

[4] N. K. Krishnan and A. S. Kiremidjian, “Time series based structural damage detection algorithm using Gausian mixtures modeling”, ***Transactions of the ASME, Journal of Dynamic system, measurement and control***, Vol. 129, pp. 285–293, 2007.

[5] ***Fundamentals of system identification***, S. Adachi, Tokyo Denki University Press (Tokyo, Japan), pp. 98-99, 2009 (in Japanese).

[6] Y. Lu and F. Gao, “A novel time-domain auto-regressive model for structural damage diagnosis”, ***Journal of Sound and Vibration***, Vol. 283, No. 3-5, pp. 1031-1049, 2005.

[7] Y. C. Zhu and S. K. Au, “Bayesian operational modal analysis with asynchronous data, part 1: Most probable value”, ***Mechanical Systems and Signal Processing***, Vol. 98, pp. 652-666, 2018.

[8] Y. C. Zhu and S. K. Au, “Bayesian operational modal analysis with asynchronous data, part 2: Posterior uncertainty”, ***Mechanical Systems and Signal Processing***, Vol. 98, pp. 920-935, 2018.

[9] M. H. Masjedian and M. Kershmiri, "A review on operational modal analysis researches: classification of methods and applications", ***In Proceedings of the 3rd International Operational Modal Analysis Conference: IOMAC*** (Portonovo, Italy) May 2009.

[10] R. Brincker, P. Anderson and L. Zhang, 'Modal identification from ambient responses using frequency domain decomposition', ***In Proceedings of the 18th International Modal Analysis Conference: IMAC*** (San Antonio, Texas, USA) February 2000.

[11] S. Takata, "Linear system identification of 1-dof vibratory system based on the Maximum Likelihood Estimation using the analytical solution of Fokker-Planck equation", ***Journal of Mechanical and Electrical Intelligent System,*** Vol. 3, No. 2, pp. 52–60, 2020.

[12] K. Araki and S. Takata, "Trial of identification of Non-linear 1-dof system based on the maximum likelihood estimation using the analytical solution of Fokker-Planck equation", ***Proceedings of Dynamics and Design Conference 2021*** (Tokyo, Japan) September 2021 (in Japanese).

[13] ***Pattern Recognition and Machine Learning (Information Science and Statistics)***, C. M. Bishop, Springer (New York, USA), 2006.

[14] B. Li and A. D. Kiureghian, "Operational modal identification using variational Bayes", ***Mechanical Systems and Signal Processing***, Vol. 88, pp. 377-398, 2017.

[15] C. Faure, J. Antoni, F. Ablitzer and C. Pezerat, "A fully Bayesian approach on Bernoulli-Gaussian prior for the identification of sparse vibratory sources from displacement measurements", ***Journal of Sound and Vibration***, Vol. 524, p. 116726, 2022.

[16] J. Daunizeau, K. J. Friston and S. J. Kiebel, "Variational Bayesian identification and prediction of stochastic nonlinear dynamic casual models", ***Physica D***, Vol. 238, pp. 2089-2118, 2009.

[17] K. Fujimoto, A. Taniguchi and Y. Nishida, "System identification of nonlinear state-space models with linearly dependent unknown parameters based on variational Bayes", ***SICE Journal of Control, Measurement, and System Integration***, Vol. 11, No. 6, pp. 456–462, 2018.

[18] A. Lund, I. Bilionis and S. J. Dyke, "Variational inference for nonlinear structural identification", ***Journal of Applied and Computational Mechanics***, Vol. 7, No. SI, pp. 1218–1231, 2021.

[19] D. Gunawan, R. Kohn and D. Nott, "Variational Bayesian approximation of factor stochastic volatility models", ***International Journal of Forecasting***, Vol. 37, pp. 1355-1375, 2021.

[20] T. K. Caughey, "Derivation and application of the Fokker-Planck equation to discrete nonlinear dynamic systems subjected to white random excitation", ***The Journal of the Acoustical Society of America***, Vol. 35, No. 11, pp. 1683–1692, 1963.