Importance Sampling for Nonlinear Oscillating Systems driven by Stationary Noise Using a Probability Measure Transformation*

Hiroaki TANAKA** and Jun OGAWA***

** Department of Applied Analysis and Complex Dynamical Systems, Graduate School of Informatics, Kyoto University, Kyoto, Japan

Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan

E-mail: hiroaki@acs.i.kyoto-u.ac.jp

*** Graduate student of Kyoto University, Kyoto, Japan

Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan

Abstract

We develop an importance sampling simulation scheme for estimating extremely small probability of level exceedance with respect to a nonlinear oscillating systems driven by a stationary random noise, where we use a probability measure transformation technique for constructing a simulation scheme based upon the Maruyama-Girsanov theorem. First, a system of Itô random differential equations and the related level exceedance are formulated, where the system nonlinearity is incorporated in terms of a nonlinear restoring force based upon the Masing rule. Next, an importance sampling simulation scheme is constructed, in which the optimal selection of the importance sampling measure is newly proposed. Finally, we give some numerical examples to demonstrate the efficiency of the proposed scheme and quantitatively examine effects of model parameters as well as the nonlinearity.

Key words : Random Differential Equation, Nonlinear Oscillating System, Nonlinear Restoring Force, Masing Rule, Importance Sampling Method, Maruyama-Girsanov Theorem

1. INTRODUCTION

Stochastic response analyses for nonlinear oscillating systems with random excitations have been playing very important roles in many fields of engineering, for example, reliability-based safety assessment for mechanical systems, a seismic response analysis against earthquakes, etc. Especially, estimation of the probability that the system response exceeds a critical value within a prespecified time interval is frequently desired from a viewpoint of system reliability assessment. It is, however, widely known that an analytical approach is generally difficult for such an estimation. Therefore, we frequently use numerical approaches such as computer simulations. The Monte Carlo method has been most widely used for such purposes.

It should be noted that convergence of the Monte Carlo scheme is generally slow. Since very small value for the exceedance probability of oscillating systems is frequently desired, such a slow convergence of the Monte Carlo simulation gives rise to too much computing efforts for generating quite many samples. It makes it actually impossible to estimate such small probability by the Monte Carlo approach. To overcome the problem, many improved methods have been developed so far. A variance reduction method has been widely used as such an improved method for reducing the variance inherent in the Monte Carlo simulation. However, most of studies reported so far are concerned with applications of variance reduction technique to time-independent systems.

In order to cope with stochastic dynamical systems, the authors have developed an importance sampling simulation scheme by applying a mathematical technique of probability

measure transformation. In Refs. (1) and (2), one of the authors has given a basic methodology for realizing an importance sampling simulation available for stochastic systems driven by a Wiener process by the use of the well-known Maruyama-Girsanov theorem^{(3),(4)} for the probability measure transformation. Further in Refs. (5) and (6), the authors have improved the method so that it can be applied to a linear oscillating system driven by a stationary noise having a prespecified power spectrum.

However, the method discussed in Refs. (5) and (6) depends closely on the linearity of the objective system. That is, we need to make further improvement to apply the method to analyze nonlinear systems. Thus, in this paper, we aim to extend the discussion in Ref. (6) to the case of a nonlinear oscillating system with random excitations. Although the general perspective has been given in Ref. (1) to construct a suitable probability measure for realizing effective reduction of computing efforts, we need to give some approximate procedure to obtain a concrete form of such probability measure. In this paper, we apply an equivalent linearization technique to construct a control function which determines the probability measure transformation.

We examine the efficiency of the proposed method for constructing an importance sampling simulation scheme through some numerical examples of estimating very small level exceedance probability associate with a nonlinear oscillating system driven by a stationary noise. The nonlinear restoring force is here supposed to obey the so-called Masing rule⁽⁷⁾ which has been widely used to describe a mathematical framework of hysteretic properties associated with nonlinear oscillating systems. Further, we quantitatively discuss on the effect of equivalent linearization having widely used in analyzing nonlinear oscillation when the level exceedance probability takes on extremely small value.

2. BASIC DESCRIPTIONS

2.1. Equation of motion

First, we give basic equations of motion describing a nonlinear oscillating system. We suppose a single degree of freedom oscillating system having the following equation of motion;

$$\ddot{x} + c\dot{x} + Q(x, \dot{x}) = F(t), \quad (1)$$

where $x(t)$ is a displacement at time t , dot represents a differentiation with respect to time t , c is a dumping coefficient, and $Q(x, \dot{x})$ corresponds to a nonlinear restoring force. The exciting force $F(t)$ is assumed to be a stochastic process, which is supposed to be a stationary process with mean zero and a prespecified power spectrum.

The nonlinear restoring force $Q(x, \dot{x})$ is here modeled using a hysteretic relation that follows the so-called Masing rule⁽⁷⁾, which has been widely used for representing the nonlinearity associating with shear deformation of metallic materials as well as reinforced concrete materials.

The characteristic of the Masing rule is represented by three basic factors: initial curve, hysteretic curve and turning point. The initial curve gives a basic nonlinear relation between the restoring force and the displacement, which is expressed as an odd function, denoted by $q_0(x)$, having a smooth property near the origin $x = 0$. Its slope at the origin represents the initial stiffness of the restoring force. Next, the hysteretic curve starting from an initial point $(x, Q) = (x_0, Q_0)$ is expressed by the use of the initial curve function $q_0(x)$ as follows;

$$Q = \begin{cases} q_0(x) & \text{(on the initial curve)} \\ Q_0 + 2q_0\left(\frac{x-x_0}{2}\right) & \text{(on the hysteretic curve).} \end{cases} \quad (2)$$

Finally, the turning point is defined as a point in the (x, Q) plain at which the velocity $\dot{x}$ changes its sign. The hysteretic curve terminates at a terminal point. If the initial point is on the initial curve, the terminal point is a symmetry point of the initial point, otherwise, the terminal point coincides with the initial point of the recent hysteretic curve.

Fig. 1 Conceptual figure of the Masing rule.

As illustrated in Fig.1, if the initial curve is defined as the curve **a**, the hysteretic curve **b** starts from the point A at which the velocity change its sign. If the hysteretic curve start from the point on the initial curve, its terminal point is decided as a symmetrical point of the initial point about the origin. So, the terminal point of the curve **b** is decided as the point B. When the velocity change its sign at the point C on the curve **b**, the hysteretic curve **c** starts from C, and its terminal point coincides with the initial point of the curve **b**, which is the initial point of the most recent hysteretic curve.

In this paper, we use the following bi-linear model for the initial curve function $q_0(x)$;

$$q_0(x) = \begin{cases} -v_0^2 u_0 + \alpha k_0(x + u_0) & (x < -u_0) \\ v_0^2 x & (|x| \leq u_0) \\ v_0^2 u_0 + \alpha k_0(x - u_0) & (x > u_0), \end{cases} \quad (3)$$

as shown in Fig.2, where u_0 , v_0 and α are yield displacement, eigen frequency in the elastic region and the so-called plastic stiffness ratio, respectively. As the ratio α decreases, the nonlinearity associated with Q becomes stronger, whereas $\alpha = 1$ corresponds to a case of linear restoring force.

2.2. Construction of a driving noise

In this paper, we use a method for constructing the driving noise $F(t)$ proposed by the authors⁽⁵⁾⁽⁶⁾, which uses an extended version of the well-known Ornstein-Uhlenbeck process⁽⁸⁾. The stationary process $F(t)$ is expressed as a superposition of plural stationary complex-valued processes, denoted by $F_1(t)$, $\dots$, $F_m(t)$, as

$$F(t) = \sum_{j=1}^m \alpha_j \text{Re}\{F_j(t)\} = \frac{1}{2} \sum_{j=1}^m \alpha_j (F_j(t) + F_j(t)^*), \quad (4)$$

where α_j are positive parameters representing weighting factors, $*$ represents a complex conjugate and $\text{Re}\{\cdot\}$ represents a real part of a complex number. Each component $F_j(t)$ ($j = 1, \dots, m$) is given as the following stochastic differential equation;

$$dF_j(t) = -\mu_j F_j(t)dt + \sigma_j dB_j(t), \quad (j = 1, 2, \dots, m), \quad (5)$$

where $B_j(t)$ ($j = 1, 2, \dots, m$) are independent standardized Wiener processes and μ_j and σ_j are

Fig. 2 An initial curve of bi-linear model.

complex parameters. By selecting μ_j and σ_j so as to satisfy

$$\sigma_j^2 = \frac{4\sigma_{0j}^2 \mu_j}{1 + \sqrt{1 + \left(\frac{\omega_{0j}}{\gamma_j}\right)^2}} \quad (j = 1, \dots, m), \quad (6)$$

with positive constants σ_{0j} ($j = 1, \dots, m$), where $\gamma_j = \text{Re}\{\mu_j\}$ and $\omega_{0j} = -\text{Im}\{\mu_j\}$ ($j = 1, \dots, m$), the power spectrum of a real part of $F(t)^j$, denoted by $S_{\text{Re}\{F_j\}}(\omega)$, is given as follows;

$$S_{\text{Re}\{F_j\}}(\omega) = \sigma_{0j}^2 \gamma_j \left(\frac{1}{\gamma_j^2 + (\omega - \omega_{0j})^2} + \frac{1}{\gamma_j^2 + (\omega + \omega_{0j})^2} \right). \quad (7)$$

Thus, the power spectrum of $F(t)$ is given as

$$S_F(\omega) = \sum_{j=1}^m \alpha_j^2 S_{\text{Re}\{F_j\}}(\omega). \quad (8)$$

Since the power spectrum given by Eq.(7) has a peak at a frequency ω_{0j} , we can construct a power spectrum of $F(t)$ which approximates a prespecified spectrum by selecting parameters γ_j and ω_{0j} ($j = 1, \dots, m$) and weighting factors α_j ($j = 1, \dots, m$)⁽⁵⁾⁽⁶⁾.

We newly introduce state variables $X_j(t)$ ($j = 1, \dots, 2m + 2$) as

$$\begin{aligned} X_1(t) &= x(t), \quad X_2(t) = \dot{x}(t), \\ X_3(t) &= F_1(t), \quad X_4(t) = F_1(t)^*, \quad \dots, \quad X_{2m+1}(t) = F_m(t), \quad X_{2m+2}(t) = F_m(t)^*, \end{aligned} \quad (9)$$

and a state vector $X(t) = (X_1(t), \dots, X_{2m+2}(t))$. Then, the response equation is given as the following system of Itô random differential equations;

$$dX_k(t) = a_k(X(t))dt + \sum_{j=1}^m b_{k,j} dB_j(t) \quad (k = 1, \dots, 2m + 2; j = 1, \dots, m), \quad (10)$$

where functions $a_k(X) = a_k(X_1, \dots, X_{2m+2})$ and coefficients $b_{k,j}$ are defined as follows;

$$\begin{aligned} a_1(X) &= X_2, \quad a_2(X) = -q(X_1, X_2) - cX_2 + \frac{1}{2} \sum_{j=1}^m \alpha_j (X_{2j+1} + X_{2j+2}), \\ a_{2k+1}(X) &= -\mu_k X_{2k+1}, \quad a_{2k+2}(X) = -\mu_k^* X_{2k+2} \quad (k = 1, \dots, m), \end{aligned} \quad (11)$$

$$b_{k,2j+1} = \begin{cases} \sigma_j & (k = j) \\ 0 & (\text{otherwise}) \end{cases}, \quad b_{k,2j+2} = \begin{cases} \sigma_j^* & (k = j) \\ 0 & (\text{otherwise}) \end{cases} \\ (j, k = 1, \dots, m). \quad (12)$$

2.3. Probability of level exceedance

Let us denote the total behavior of the solution process $X(t)$ of Eq.(1) as $X \equiv \{X(t); 0 \leq t \leq T\}$. Each sample of X is an element of a function space $C([0, T] : \mathbb{R}^{2m+2})$, which is a set of all $\mathbb{R}^{2m+2}$ -valued continuous functions defined in $[0, T]$. The probability of level exceedance up to time T , denoted by $p_f(T)$, is then expressed as follows;

$$p_f(T) = E_P \{f[X]\}, \quad (13)$$

where $E_P\{\cdot\}$ is an operator to take expectation under the original probability measure P and $f : C([0, T] : \mathbb{R}^{2m+2}) \rightarrow \mathbb{R}^1$ is a functional defined as,

$$f[X] = \begin{cases} 1 & (0 \leq t \leq T \text{ s.t. } L(X(t)) \leq 0) \\ 0 & (\text{otherwise}) \end{cases}. \quad (14)$$

The function $L : \mathbb{R}^{2m+2} \rightarrow \mathbb{R}^1$ is the so-called limit state function, where $\{X; L(X) \leq 0\}$ expresses the region corresponding to the level exceedance. Here we basically pay attention to the level exceedance of the displacement $X_1(t)$, i.e.,

$$L(X) = \min \{L^{(+)}(X), L^{(-)}(X)\}, \quad (15)$$

$$L^{(+)}(X) = x_c - X_1, \quad L^{(-)}(X) = x_c + X_1, \quad (16)$$

where x_c is a critical displacement. If the system falls into a *failure* state when the system state exceeds such a critical level, $p_f(T)$ gives the probability that the system fails within time T , which is frequently called *probability of failure*. Hereinafter in this paper, we call $p_f(T)$ *probability of failure*, which is a kind of generalized term including the probability of simple level exceedance as a special case.

Let us denote B as the total behavior of the noise process similarly as X . The total behavior X is related to B through Eq.(1). We formally express it as $X = \mathcal{H}[B]$ for each sample, where the mapping $\mathcal{H} : C([0, T] : \mathbb{R}^m) \rightarrow C([0, T] : \mathbb{R}^{2m+2})$ is determined by Eq.(1). Using the map $\mathcal{H}$, we can rewrite Eq.(13) as

$$p_f(T) = E_P \{f[\mathcal{H}[B]]\}. \quad (17)$$

3. MONTE CARLO SIMULATION SCHEME AND ITS IMPROVEMENT

3.1. Monte Carlo scheme based upon a probability measure transformation

It is generally well-known that we can not hardly, except for a few simple cases, give an analytical expression for the probability of failure $p_f(T)$ given by Eq.(17). Thus, we need to take a numerical approach for estimating it, such as the well-known Monte Carlo method. However, since extremely small probability of failure is desired for important systems, the crude Monte Carlo method based upon Eq.(17) requires too many samples.

To avoid such a difficulty we apply a method based upon a probability measure transformation as discussed in Ref. (6). Introducing another probability measure $\tilde{P}$ that is assumed to be equivalent with the original probability measure P , we rewrite Eq.(17) as

$$p_f(T) = E_{\tilde{P}} \left\{ f[\mathcal{H}[B]] \frac{dP}{d\tilde{P}} \right\} \quad (18)$$

where $E_{\tilde{P}}\{\cdot\}$ is an operator to take expectation under the newly introduced measure $\tilde{P}$ and $dP/d\tilde{P}$ expresses the Radon-Nikodym derivative between these two measures. A Monte Carlo estimator corresponding to Eq.(18) is then given as

$$\hat{p}_f(T; N) = \frac{1}{N} \sum_{j=1}^N f[\mathcal{H}[B_{\tilde{P}}^{(j)}]] \left(\frac{dP}{d\tilde{P}} \right)_{\tilde{P}}^{(j)}, \quad (19)$$

where $B_P^{(j)}$ ($j = 1, 2, \dots$) and $(dP/d\tilde{P})_P^{(j)}$ ($j = 1, 2, \dots$) are independent simulations of B and $dP/d\tilde{P}$ respectively generated according to the probability measure $\tilde{P}$.

We can easily show that the Monte Carlo estimator given by Eq.(19) is unbiased and consistent, i.e., $\hat{p}_f(T; N)$ converges to the target value $p_f(T)$ in the sense of mean square convergence as $N \rightarrow \infty$. Further, it is well-known that the variance $\hat{p}_f(T; N)$ inherent in the Monte Carlo procedure can be reduced by selecting a suitable probability measure for $\tilde{P}$ so that we can generate many samples of failure. Thus, the measure $\tilde{P}$ is frequently called an importance sampling measure⁽²⁾.

3.2. Application of the Maruyama-Girsanov theorem

To execute the importance sampling simulation given by Eq.(19), we need to select the importance sampling measure $\tilde{P}$ so that (i) we can easily generate samples of the Wiener process B under $\tilde{P}$ and (ii) we can give an analytical expression for the Radon-Nikodym derivative $dP/d\tilde{P}$. Thus, we have to confine ourselves to a rather restricted transformation of the probability measure.

One of the most effective approaches uses the well-known Maruyama-Girsanov transformation^{(3),(4)}. Let us define, using the Wiener process $B(t)$, an $\mathbb{R}^m$ -valued random process $\tilde{B}(t) = (\tilde{B}_1(t), \dots, \tilde{B}_m(t))$ as

$$\tilde{B}_j(t) = B_j(t) - \int_0^t u_j(s)ds \quad (j = 1, \dots, m), \quad (20)$$

where $u(t) = (u_1(t), \dots, u_m(t))$ is generally an $\mathbb{R}^m$ -valued stochastic process having a suitably integrable property. However, we suppose that $u(t)$ is an $\mathbb{R}^m$ -valued deterministic function with a suitably integrable property as discussed by Tanaka⁽¹⁾. Although the process $\tilde{B}(t)$ is not a Wiener process under the probability measure P , the Maruyama-Girsanov theorem^{(3),(4)} shows that it is a Wiener process under a probability measure $\tilde{P}$, which is absolutely continuous with respect to the original probability measure P , whose Radon-Nikodym derivative is given as follows;

$$\frac{dP}{d\tilde{P}} = \exp \left[\sum_{j=1}^m \left\{ - \int_0^T u_j(t)dB_j(t) + \frac{1}{2} \int_0^T u_j(t)^2 dt \right\} \right] \quad (21)$$

Substituting Eq.(20) into Eq.(10), we can obtain the following system of Itô random differential equations;

$$dX_k(t) = \left\{ a_k(X(t)) + \sum_{j=1}^m b_{k,j}u_j(t) \right\} dt + \sum_{j=1}^m b_{k,j}d\tilde{B}_j(t) \quad (k = 1, \dots, 2m+2), \quad (22)$$

Since $\tilde{B}(t)$ is a Wiener process under $\tilde{P}$, the mean behavior of the solution $X(t)$ can be controlled by giving suitable behavior for $u(t)$. Consequently, we can predominantly generate many samples arriving at the failure region, which leads to the variance reduction in estimating the quite small probability of failure.

4. SELECTION OF THE SAMPLING MEASURE $\tilde{P}$

4.1. Basic framework for selecting sampling control function

In Ref. (1), one of the authors has discussed a systematic approach for deriving the optimal sampling control function by the use of a concept of 'design point'⁽⁹⁾ in the so-called first-order second-moment approach⁽¹⁰⁾. To this end, introducing a difference scheme $t_k = k\Delta t$ ($k = 1, \dots, K$; $\Delta t = T/K$), we define a family of random variables as

$$\Xi_{j,k} \equiv \frac{B_j(t_k) - B_j(t_{k-1})}{\sqrt{\Delta t}} \quad (j = 1, \dots, m; k = 1, \dots, K). \quad (23)$$

According to the basic property of the Wiener process, $\Xi_{1,1}, \dots, \Xi_{m,K}$ are mutually independent standardized Gaussian random variables. The vector space spanned by $\Xi = (\Xi_{1,1}, \dots, \Xi_{m,K})$ is frequently called a *standardized space*. The solution $x(t) = X_1(t)$ at time t_k can be approximately expressed by the use of Ξ as $X_1(t_k) \simeq \mathcal{Y}_k(\Xi)$, where the function $\mathcal{Y}_k$ is determined

by applying a suitable difference scheme for Eq.(10). The boundary of the failure zone corresponding to the limit state function $L^{(+)}(X)$ is given as

$$\mathcal{Y}_k(\Xi) = x_c \quad (k = 1, \dots, K), \quad (24)$$

at time t_k , which gives a surface in Ξ -space. The so-called design point is defined as a point on the surface given by Eq.(24) that minimizes a distance from the origin of Ξ -space. That is, the design point at time t_k in the standardized space is obtained by solving the following minimization problem;

$$\text{minimize } \sum_{k=1}^K \sum_{j=1}^m (\Xi_{j,k})^2, \quad \text{subject to } \mathcal{Y}_k(\Xi) = x_c \quad (k = 1, \dots, K). \quad (25)$$

The distance between the design point and the origin, denoted by β_k , represents a safe margin at time t_k . Hence, the time corresponding to the minimum of $\{\beta_1, \dots, \beta_K\}$ plays the most important role for evaluating the probability of failure. Such time is called *design time*⁽¹⁾. Let t_{k_d} be a design time and $\hat{\Xi}_{j,k}(t_{k_d})$ ($k = 0, \dots, K$; $j = 1, \dots, m$) be a design point obtained at time t_{k_d} . It has been well-known that the effective convergence can be realized by locating the mean of the system state under $\tilde{P}$ on the design point⁽⁹⁾. Since the increment $B_j(t_k) - B_j(t_{k-1})$ obeys, according to the Girsanov transformation given by Eq.(20), a Gaussian distribution with mean $u_j(t_{k-1})\Delta t$ under $\tilde{P}$, such a situation can be realized by selecting $u_j(t)$ as satisfying the relation,

$$u_m(t_{k-1})\sqrt{\Delta t} = \hat{\Xi}_{j,k}(t_{k_d}) \quad (k = 1, \dots, K; j = 1, \dots, m). \quad (26)$$

4.2. Approximation by an equivalent linearization

If the basic equation of motion is linear, the approximating function $\mathcal{Y}_k$ appearing in Eq.(24) becomes linear. Thus, the solution of the minimization problem given by Eq.(25) can be obtained in a closed form as shown in Ref. (6). However, since it is generally difficult to obtain such a solution when the equation of motion is nonlinear, we need to take one of the following approaches; linear approximation of the equation of motion, linear approximation of $\mathcal{Y}_k(\Xi)$ and transformation of the dependent variables⁽²⁾.

Among them, a transformation of the dependent variables is not applicable to our basic model. Further, a linear approximation of $\mathcal{Y}_k(\Xi)$ is not easy since we need to give an explicit form of $\mathcal{Y}_k(\Xi)$. Therefore, in this paper, we apply a linear approximation of the equation of motion. Supposing a sinusoidal input acceleration with a frequency ν_0 , which is an eigen frequency in the elastic region appearing in Eq.(3), we consider an equation motion as

$$\ddot{x} + c\dot{x} + Q(x, \dot{x}) = A \sin \nu_0 t, \quad (27)$$

where we set the amplitude A as the magnitude of the critical response x_c so that the mean behavior under the importance sampling measure arrives at the failure surface. To obtain an equivalent linearized equation, we use the so-called least square approximation⁽¹¹⁾ with a linear equation with error term as follows ;

$$\ddot{x} + c_{\text{eff}}\dot{x} + \omega_{\text{eff}}^2 x + \varepsilon(x, \dot{x}, t) = A \sin \nu_0 t, \quad (28)$$

where c_{eff} and ω_{eff} are effective damping coefficient and effective eigen frequency respectively. The effective coefficients c_{eff} and ω_{eff} are obtained by minimizing the mean of squared error term ε ;

$$\overline{\varepsilon^2} = \overline{\{(c_{\text{eff}} - c)\dot{x} + \omega_{\text{eff}}^2 x - Q(x, \dot{x})\}^2}. \quad (29)$$

well approximates the stationary response of Eq.(27). In Eq.(29), the upper bar represents a mean taken over one period when the solution shows almost stationary behavior.

Based upon Eq.(28), we consider a system of Itô random differential equations obtained by replacing the basic equation of motion Eq.(1) by

$$\ddot{x} + c_{\text{eff}}\dot{x} + \omega_{\text{eff}}^2 x = F(t). \quad (30)$$

Provided that the linearized version given by Eq.(29) can well reproduce the mean behavior of the solution process, we can apply our result in Ref. (6) for constructing the sampling control function. That is, the design time is equal to the time horizon T and the optimal sampling control function corresponding to a limit state function $L^{(+)}(X)$, denoted by $\hat{u}(t) = (\hat{u}_1(t), \dots, \hat{u}_m(t))$, is given as follows;

$$\hat{u}_j(t) = \frac{\text{Re}\{\sigma_j \varphi_j(T-t)\}}{\sum_{j'=1}^m \int_0^T \text{Re}\{\sigma_{j'} \varphi_{j'}(T-t)\}^2 dt} x_c \quad (0 \leq t \leq T; j = 1, \dots, m) \quad (31)$$

$$\varphi_j(t) = \frac{1}{\kappa_j} \left\{ e^{-\mu_j t} - \frac{\mu_j + \lambda_2}{\lambda_2 - \lambda_1} e^{\lambda_1 t} + \frac{\mu_j + \lambda_1}{\lambda_2 - \lambda_1} e^{\lambda_2 t} \right\} \quad (j = 1, \dots, m) \quad (32)$$

$$\kappa_j = \frac{2}{\alpha_j} \left\{ (-\mu_j)^2 + c_{\text{eff}} \omega(-\mu_j) + \omega_{\text{eff}}^2 \right\} \quad (j = 1, \dots, m) \quad (33)$$

where λ_1 and λ_2 are solutions of the quadratic equation $\lambda^2 + c_{\text{eff}} \lambda + \omega_{\text{eff}}^2 = 0$. The optimal control function corresponding to a limit state function $L^{(-)}(X)$ can be easily obtained by changing x_c by $-x_c$ in Eq.(31).

It should be noted that the linear approximation is only applied to construct a sampling control function, i.e., our Monte Carlo simulation is executed based upon a system of nonlinear equations given by Eq.(10).

As for an application of an equivalent linearization procedure to an importance sampling simulation, Olsen and Naess⁽¹³⁾ have proposed to apply the linearization procedure to an iteration procedure for constructing a sampling measure based upon the so-called adaptive sampling principle⁽¹⁴⁾. Therefore, our approach is intrinsically different from Olsen and Naess. It should be emphasized that our approach does not need no *pre*-simulation efforts for constructing an importance sampling measure.

4.3. Importance sampling measure with plural design times

Under the importance sampling measure corresponding to a specified design time, most of samples of $X_1(t)$ arrive at the limit state surface near the specified design time. Such constructed importance sampling measure gives a good performance for evaluating very small probability of failure, provided that the “failure mode” is dominant, for example, as shown in Ref. (2). However, since the system response discussed here tends to almost stationary behavior as the time horizon T becomes large, which degrades the accuracy of the estimation under the importance sampling scheme based upon one design time, since other many failure modes contribute to the probability of system failure in the specified time interval $[0, T]$.

Thus, plural design points are needed for constructing an importance sampling measure⁽¹²⁾. When the time horizon T is large enough so that the solution $X_1(t)$ shows almost stationary behavior in the whole region of $[0, T]$, design times corresponding to the limit state function $L^{(+)}$ (resp. $L^{(-)}$), denoted by $t_{d(+)}^\ell$ ($\ell = 1, \dots, J$) (resp. $t_{d(-)}^\ell$ ($\ell = 1, \dots, J$)), are located with equi-intervals as

$$t_{d(+)}^\ell = t_{d(-)}^\ell = \frac{T}{J} \ell \quad (\ell = 1, \dots, J). \quad (34)$$

The optimal sampling control function $\hat{u}_j(t)$ corresponding to the design time $t_{d(+)}^\ell$ is given as

$$\hat{u}_j(t) = \begin{cases} \frac{\text{Re}\{\sigma_j \varphi_j(t_{d(+)}^\ell - t)\}}{\sum_{j'=1}^m \int_0^{t_{d(+)}^\ell} \text{Re}\{\sigma_{j'} \varphi_{j'}(t_{d(+)}^\ell - t)\}^2 dt} x_c & (t \leq t_{d(+)}^\ell) \\ 0 & (t > t_{d(+)}^\ell) \end{cases} \quad (j = 1, \dots, m). \quad (35)$$

The control function $\hat{u}_j(t)$ corresponding to the design time $t_{d(-)}^\ell$ can be easily obtained by replacing x_c in Eq.(35) by $-x_c$.

The importance sampling measure $\tilde{P}$ is then constructed as follows;

$$\tilde{P} = \frac{1}{2J} \sum_{\ell=1}^J (\tilde{P}_{\ell}^{(+)} + \tilde{P}_{\ell}^{(-)}), \quad (36)$$

where $\tilde{P}_{\ell}^{(+)}$ (resp. $\tilde{P}_{\ell}^{(-)}$) represents an importance sampling measure corresponding to the sampling control function with a design point $t_{d(+)}^{\ell}$ (resp. $t_{d(-)}^{\ell}$) ($\ell = 1, \dots, J$) given by Eq.(35). Here we suppose same weighting factor for each importance sampling measure $\tilde{P}_{\ell}^{(+)}$ and $\tilde{P}_{\ell}^{(-)}$, provided that the time horizon T is large enough. Therefore, the Radon-Nikodym derivative can be finally constructed as follows;

$$\frac{dP}{d\tilde{P}} = \left[\frac{1}{2J} \sum_{\ell=1}^J \left\{ \left(\frac{dP}{d\tilde{P}_{\ell}^{(+)}} \right)^{-1} + \left(\frac{dP}{d\tilde{P}_{\ell}^{(-)}} \right)^{-1} \right\} \right]^{-1} \quad (37)$$

5. NUMERICAL EXAMPLES

In this section, we give some numerical examples to demonstrate the efficiency of our proposed simulation scheme, and examine some parameter sensitivities as well as the effect of system nonlinearity.

5.1. Numerical procedure based upon the Masing rule

Here we briefly summarize a numerical integration procedure for Eq.(10). Let $t_i = i\Delta t$ ($i = 0, 1, 2, \dots$) be i -th mesh point of the time variable and we use simple notations such as $\Delta B_j(t_i) \equiv B_j(t_i) - B_j(t_{i-1})$. Applying the Heun scheme, we can give a discretized version of Eq.(10) as follows;

$$X_k(t_{i+1}) = X_k(t_i) + \frac{1}{2} \{a_k(Y(t_i)) + a_k(X(t_i))\} \Delta t + \sum_{j=1}^m b_{k,j} \Delta B_j(t_i) \quad (k = 1, \dots, 2m+2), \quad (38)$$

with a supporting value

$$Y_k(t_i) = X_k(t_i) + a_k(X(t_i))\Delta t + \sum_{j=1}^m b_{k,j} \Delta B_j(t_i) \quad (k = 1, \dots, 2m+2). \quad (39)$$

Since the drift coefficient $a_2(X)$ contains the restoring term $Q(X_1, X_2)$, which follows the Masing rule⁽⁷⁾, the value of Q at the i -th mesh point is determined by the displacement and velocity on the previous two mesh points t_i and t_{i-1} . If the velocity $X_2(t)$ changes its sign between t_i and t_{i-1} , the turning point, where the velocity becomes zero, is set as the intermediate point of the interval. The turning points appear every time the velocity changes its sign. The displacement X_1^{TP} and the restoring term Q^{TP} at the turning point is determined by a linear interpolation as follows;

$$X_1^{TP} = \frac{X_2(t_i)X_1(t_{i-1}) - X_2(t_{i-1})X_1(t_i)}{X_2(t_i) - X_2(t_{i-1})}, \quad (40)$$

$$Q^{TP} = Q(X_0, 0). \quad (41)$$

Then, once $|X_1|$ exceeds u_0 , $Q(t_i)$ is given as;

$$Q(t_i) = Q_0 + 2q_0 \left(\frac{X_1(t_i) - X_{1,0}}{2} \right), \quad (42)$$

where q_0 is defined as Eq.(3).

5.2. Results of simulations and discussions

Next, we show results of simulations, where parameters in the basic oscillating system are selected as: the eigen frequency in the elastic region $\nu_0 = 10.0[\text{s}^{-1}]$ and the dumping ratio $c = 0.03$. Parameters associated with the power spectrum of the input noise $F(t)$ are selected as

Fig. 3 Several samples of the input noise $F(t)$.

Fig. 4 Several samples of the system response $X_1(t)$.

$$m = 1, \alpha_1 = 1, \omega_{01} = \pi[\text{s}^{-1}], \gamma_1 = 1.8[\text{s}^{-1}], \sqrt{\sigma_{01}/u_0} = 5.0[\text{s}^{-1}], \quad (43)$$

i.e., we suppose a spectrum having one peak at a frequency $\omega_{01} = 3.0[\text{s}^{-1}]$. It should be noted that we can easily treat other spectrum of more general form (6). In each simulations, the time horizon in the simulation T is set as 10.0 [s]. Figures 3 and 4 give several samples of the input noise $F(t)$ and the response $X_1(t)$ under parameters set as above. We can see that the response $X_1(t)$ shows almost stationary variation. Thus, we need to apply a scheme with plural design times as discussed in Section 4.3. In this paper, we arrange design times with constant intervals slightly less than the correlation time of the input noise $F(t)$.

5.2.1. Estimated probability of failure Figures 5 and 6 show results of simulations, where the estimated probability of failure is logarithmically plotted for several values of a normalized critical displacement x_c/u_0 . In both figures, open diamonds indicate results of the crude Monte Carlo simulation with 10^6 samples, and open squares indicate results of the proposed method with 2000 samples and each error bar shows a width of standard deviation of the estimator obtained by 20 times independent simulations. The results show that our proposed scheme can estimate very small probability of failure (less than 10^{-6}) with small sample size ($N = 2000$), even if the nonlinearity is strong ($\alpha = 0.2$). On the other hand, if the desired probability of failure is very small, the crude Monte Carlo method needs very large sample size ($N = 10^6$) and thus very large computing time (about 300 times of the proposed scheme). Although with such efforts, the estimated probability obtained by the crude Monte Carlo method shows large variance, i.e., worse convergence in the simulation. Especially if x_c/u_0 is greater than 3.5, some of estimations by the crude Monte Carlo method become zero.

To figure out the difference of the convergence between the both simulation more clearly, we execute the following comparison. Figure 7 shows the coefficient of variation of the estimated probability of failure in the crude Monte Carlo simulation and the proposed scheme with the same numbers of samples ($N = 5000$) for several values of x_c/u_0 , where the coefficient of variation is calculated through 20 times independent simulations. We can see that our proposed scheme gives very good convergence regardless of the critical value x_c/u_0 , whereas the convergence of the crude Monte Carlo method drastically deteriorates as x_c/u_0 increases, i.e., the target probability becomes small.

Fig. 5 Estimated probability of failure in the case of weak nonlinearity(\$\alpha = 0.8\$).

Fig. 6 Estimated probability of failure in the case of strong nonlinearity(\$\alpha = 0.2\$).

Fig. 7 Comparison of the coefficient of variation in the crude Monte Carlo simulation (CMC) and the proposed scheme (FMC).

Fig. 8 Sample spectra for several values of $\gamma = \text{Re}\{\mu\}$.

Fig. 9 Effect of a band width of the input noise on the coefficient of variation.

5.2.2. Effect of a band width of an input spectrum on the convergence of the proposed simulation scheme

Here, we examine the effect of the spectrum of the input noise $F(t)$ on the convergence of the proposed simulation scheme. Figure 8 shows the power spectrum $S_F(\omega)$ of the input noise $F(t)$ for several value of γ_1 . We can see that the band width of spectrum narrows as γ_1 becomes small. Figure 9 shows the effect of the parameter γ_1 , i.e., the effect of the band width on the coefficient of variation of the estimated probability of failure. Although the coefficient of variation becomes slightly large when γ_1 takes on a very small value, we can see that almost good convergence is achieved regardless of the band width of the power spectrum of the input noise.

5.2.3. Effect of the linearization in the region of small probability of failure

Next, we examine the effect of the equivalent linearization on the probability of failure. In Figs.10 and 11, we plot the estimated probability of failure based upon the nonlinear model given by Eq.(1) by the use of our proposed scheme with 2000 samples (open squares) and that based upon the linearized model given by Eq.(30) by the use of crude Monte Carlo method with 10^6 samples (solid crosses) for several values of critical displacement, where the nonlinearity of the restoring force is weak ($\alpha = 0.8$). On the other hand, results in the case of strong nonlinearity ($\alpha = 0.2$) are shown.

When the nonlinearity is weak ($\alpha = 0.8$), the equivalent linearized model can give a good approximation for the probability of failure even if it takes on very small values. However, when the nonlinearity is strong ($\alpha = 0.2$), the equivalent linearized model gives underesti-

Fig. 10 Comparison of the estimated probability of failure with results obtained by the equivalent linear model in the case of weak nonlinearity ($\alpha = 0.8$).

Fig. 11 Comparison of the estimated probability of failure with results obtained by the equivalent linear model in the case of strong nonlinearity ($\alpha = 0.2$).

mation for the probability of failure especially in the region of extremely small probability of failure. Thus, the equivalent linearized model may give non-conservative estimation, which is critically important in practical seismic safety analysis, i.e., the linearized model is not appropriate for estimating extremely small probability of failure when the system shows strong nonlinearity.

It should be emphasized that such property can be quantitatively clarified only by applying our proposed simulation scheme for estimating extremely small probability of failure.

6. CONCLUSIONS

In this paper, we have improved the importance sampling method proposed by one of the authors⁽¹⁾, so that we can treat stochastically oscillating systems having nonlinear restoring force. Based upon the equivalent linearization procedure, we have proposed a systematic procedure to construct the optimal sampling control function, under which we can most effectively reduce the variance inherent in the Monte Carlo simulation. Consequently we can estimate extremely small probability of failure of the nonlinear oscillating system excited by a stationary random noise.

We have shown some numerical examples to demonstrate the effectiveness of our proposed simulation scheme. The results show that the proposed scheme works quite well for estimating extremely small probability of failure even if the nonlinearity, associated with the restoring force, is strong. We have also shown that the linearly approximated system gives rather small value for the probability of failure compared with the results obtained from the

nonlinear system in the region of small probability of failure. Such a result tells us that the equivalent linearization is not appropriate for quantifying the safety of oscillating systems. It should be emphasized that such a property is clarified only by applying our improved scheme.

In our approach, we construct a sampling control function by the use of an equivalent linearized model. Although such an approach is expected to be applicable to nonlinear oscillating systems of more general type, rigorous verification of such a wide applicability has not been given in this paper, which is still an important open problem.

In future, we need to give a further discussions to construct an importance sampling measure so that we can effectively reduce the variance in the simulation and to deal with a non-linearity of limit surface. We also need to discuss on the application of the proposed method to more general systems. For example, we consider that the extension to large dimensional oscillators can be easily derived, if the system is linear.

References

- (1) H. Tanaka, An importance sampling simulation scheme for time-dependent system reliability analyses using the Girsanov transformation, *Proc. of ICOSSAR'97* 1 (1998) 411–418.
- (2) H. Tanaka, Importance sampling simulation for a stochastic fatigue crack growth model, *Proc. of ICASP8**1999 2 (1999) 907–914.
- (3) G. Maruyama, Continuous Markov processes and stochastic equations, *Rend. Circ. Matem. Palermo* 10 (1955) 48–90.
- (4) I. V. Girsanov, On transforming a certain class of stochastic processes by absolutely continuous substitution, *Theory Probab. Its Appl.* 5 (1960) 285–301.
- (5) J. Ogawa, H. Tanaka, Fast Monte Carlo simulation for stochastic response analysis under stationary noise having a specified power spectrum, *Applications of Statistics and Probability in Civil Engineering (Proc. of ICASP 10)*, CD-ROM paper No.029, Taylor & Francis/Balkema (2007)
- (6) J. Ogawa, H. Tanaka, Importance Sampling for Stochastic Systems under Stationary Noise Having a Specified Power Spectrum, *Probabilistic Engineering Mechanics*, 24 (2009), 537–544.
- (7) G. Masing, Eigenspannungen und Verfestigung beim Messing, *Proc. Second Internatl. Congr. App. Mechs.* (1926).
- (8) G. E. Uhlenbeck, L. S. Ornstein, On the theory of Brownian motion, *Physical Review* 36 (1930) 823–841.
- (9) U. Bourgund, C. G. Bucher, Importance Sampling Procedure Using Design point – ISPUD – A User's Manual, Institute of Engineering Mechanics, University of Innsbruck, 1986.
- (10) See, for example, K. Dolinski, First-order second-moment approximation in reliability of structural systems: Critical review and alternative approach, *Structural Safety*, 1 (1983), 211–231.
- (11) T. K. Caughey, Equivalent Linearization Techniques, *Journal of the Acoustical Society of America*, Vol.90, No. EM2 (1963) 1706–1711.
- (12) H. Tanaka, A fast monte Carlo method for stochastic systems showing stationary behavior, *Proc. of JCOSSAR'07* 6 (2007) 111–116 (in Japanese).
- (13) A. Ivanova Olsen, A. Naess, An importance sampling procedure for estimating failure probabilities of non-linear dynamic systems subjected to random noise, *International Journal of Non-Linear Mechanics* 42 (2007) 848–863.
- (14) C. G. Bucher, Adaptive Sampling - An Iterative Fast Monte Carlo Procedure, *Strutural Safety*, 5 (1988) 119–126.