SIAM J. NUMER. ANAL.
Vol. 59, No. 1, pp. 32–59

ASYMPTOTICALLY-PRESERVING LARGE DEVIATIONS PRINCIPLES BY STOCHASTIC SYMPLECTIC METHODS FOR A LINEAR STOCHASTIC OSCILLATOR*

CHUCHU CHEN†, JIALING HONG†, DIANCONG JIN‡, AND LIYING SUN†

Abstract. It is well known that symplectic methods have been rigorously shown to be superior to nonsymplectic ones especially in long-time computation, when applied to deterministic Hamiltonian systems. In this paper, we attempt to study the superiority of stochastic symplectic methods by means of the large deviations principle. We propose the concept of asymptotical preservation of numerical methods for large deviations principles associated with the exact solutions of the general stochastic Hamiltonian systems. Considering that the linear stochastic oscillator is one of the typical stochastic Hamiltonian systems, we take it as the test equation in this paper to obtain precise results about the rate functions of large deviations principles for both exact and numerical solutions. Based on the Gärtner–Ellis theorem, we first study the large deviations principles of the mean position and the mean velocity for both the exact solution and its numerical approximations. Then, we prove that stochastic symplectic methods asymptotically preserve these two large deviations principles, but nonsymplectic ones do not. This indicates that stochastic symplectic methods are able to approximate well the exponential decay speed of the "hitting probability" of the mean position and mean velocity of the stochastic oscillator. Finally, numerical experiments are performed to show the superiority of stochastic symplectic methods in computing the large deviations rate functions. To the best of our knowledge, this is the first result about applying the large deviations principle to reveal the superiority of stochastic symplectic methods compared with nonsymplectic ones in the existing literature.

Key words. stochastic symplectic methods, superiority, large deviations principle, rate function, asymptotical preservation

AMS subject classifications. 60F10, 60H35, 65C30, 65P10

DOI. 10.1137/19M1306919

1. Introduction.

A $2d$-dimensional stochastic differential equation (SDE) is called a stochastic Hamiltonian system if it can be written in the form

$$\mathrm{d}\begin{pmatrix} p \\ q \end{pmatrix} = J^{-1}\nabla H_0(p,q)\mathrm{d}t + \sum_{r=1}^{m} J^{-1}\nabla H_r(p,q)\circ \mathrm{d}W_r(t), \qquad J = \begin{bmatrix} 0 & I_d \\ -I_d & 0 \end{bmatrix}, \tag{1.1}$$

where $\circ$ denotes the Stratonovich product, H_i, $i=0,1,\dots,m$ are smooth Hamilton functions, and $W=(W_1,\dots,W_m)$ is an m-dimensional Brownian motion on a given complete filtered probability space $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geq 0},\mathbf{P})$. The phase flow of (1.1) preserves the symplectic structure in phase space, i.e., $dp(t)\wedge dq(t) = dp(0)\wedge dq(0)$, a.s., for all $t\geq 0$. In order to preserve the symplectic structure, a class of numerical methods called stochastic symplectic methods are proposed [18]. In recent years, stochastic symplectic methods have received extensive attention, and large quantities

*Received by the editors December 16, 2019; accepted for publication (in revised form) September 28, 2020; published electronically January 4, 2021.
https://doi.org/10.1137/19M1306919
Funding: The work of the authors was supported by the National Natural Science Foundation of China grants 91630312, 11711530071, 11871068, 11971470, 12031020, and 12022118.
†LSEC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, and School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China (chenchuchu@lsec.cc.ac.cn, hjl@lsec.cc.ac.cn, liyingsun@lsec.cc.ac.cn).
‡Corresponding author. LSEC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, and School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China (diancongjin@lsec.cc.ac.cn).

of numerical experiments show that stochastic symplectic methods possess excellent long-time stability (see, e.g., [3, 5, 6, 9, 12, 15, 23, 24]). One approach to theoretically explaining the superiority of stochastic symplectic methods is based on the techniques of modified equations and backward error analysis (see [1, 2, 4, 10, 14, 22, 24, 25] and references therein). Different from their approach, we try to apply the large deviations principle (LDP) to discover the superiority of the stochastic symplectic methods in this paper.

The large deviations principle is concerned with the exponential decay of probabilities of rare events, which can be regarded as an extension or refinement of the law of large numbers and central limit theorem. It is usually used to describe the asymptotical behavior of stochastic processes for which large deviations estimates are concerned. If a stochastic process $\{X_T\}_{T>0}$ satisfies an LDP with the rate function I, then the *hitting probability* $\mathbf{P}\left(X_T \in [a, a+\mathrm{d}a]\right)$ decays exponentially, i.e., $e^{-TI(a)}\mathrm{d}a$. The rate function characterizes the fluctuations of the stochastic process $\{X_T\}_{T>0}$ in the long-time limit and has a wide range of applications in engineering and physical sciences (see, e.g., [13]). When a numerical method is applied to a given stochastic differential equation, it is worthwhile to study whether the numerical method can preserve asymptotically the decay rate e^{-TI}.

Let $\{Z_T\}_{T>0}$ be a stochastic process associated with the exact solution of (1.1), usually viewed as an observable of (1.1). For a numerical method $\{p_n, q_n\}_{n\geq 0}$ approximating (1.1), let $\{Z_{T,N}\}_{N\geq 1}$ be a discrete approximation of $\{Z_T\}_{T>0}$ associated with the numerical method $\{p_n, q_n\}_{n\geq 0}$. For example, one can take $Z_T = \frac{1}{T}\int_0^T f(p(t), q(t))\mathrm{d}t$ as an observable of (1.1) for some smooth function f. Then $Z_{T,N} = \frac{1}{N}\sum_{n=0}^{N-1} f(p_n, q_n)$ can be viewed as a discrete version of Z_T. If $\{Z_T\}_{T>0}$ satisfies an LDP with the rate function I, one natural question is the following:

(Q1) Does $\{Z_{T,N}\}_{N\geq 1}$ satisfy the LDP with some rate function I^h for a fixed step-size h?

If so, then $\{Z_T\}_{T>0}$ and $\{Z_{T,N}\}_{N\geq 1}$ formally satisfy

$$\mathbf{P}\left(Z_T \in [a, a+\mathrm{d}a]\right) \approx e^{-TI(a)}\mathrm{d}a \text{ for sufficiently large } T, \tag{1.2}$$

$$\mathbf{P}(Z_{T,N} \in [a, a+\mathrm{d}a]) \approx e^{-NI^h(a)}\mathrm{d}a = e^{-t_N I^h(a)/h}\mathrm{d}a \quad \text{for sufficiently large } t_N. \tag{1.3}$$

With $T = t_N$ being the observation scale, it is reasonable to use I^h/h to evaluate the ability of the numerical method to preserve the large deviations rate function. Hence, another meaningful question is the following:

(Q2) If $\{Z_{T,N}\}_{N\geq 1}$ satisfies an LDP with the rate function I^h, could I^h/h approximate I well for sufficiently small h?

Concerning the above questions, we give the following definition on the asymptotical or even exact preservation of LDP.

Definition 1.1. *Let E be a Polish space, i.e., complete and separable metric space. Let $\{Z_T\}_{T>0}$ be a stochastic process associated with the exact solution of* (1.1). *Let $\{Z_{T,N}\}_{N\geq 1}$ be a discrete approximation of $\{Z_T\}_{T>0}$, associated with some numerical method $\{p_n, q_n\}_{n\geq 0}$ for* (1.1). *Assume that $\{Z_T\}_{T>0}$ and $\{Z_{T,N}\}_{N\geq 1}$ satisfy the LDPs on E with the rate functions I and I^h, respectively. We call $I^h_{mod} := I^h/h$ the modified rate function of I^h. Moreover, the numerical method $\{p_n, q_n\}_{n\geq 0}$ is said to asymptotically preserve the LDP of $\{Z_T\}_{T>0}$ if*

$$\lim_{h\to 0} I^h_{mod}(y) = I(y) \qquad \forall \quad y \in E. \tag{1.4}$$

In particular, the numerical method $\{p_n, q_n\}_{n\geq 0}$ is said to exactly preserve the LDP of $\{Z_T\}_{T>0}$ if for all sufficiently small step-size h, $I^h_{mod}(\cdot) = I(\cdot)$.

Concerning that the linear stochastic oscillator is one of the typical stochastic Hamiltonian systems, we take it as the test equation in this paper to obtain precise results about the rate functions of LDPs for both the exact solution and a class of numerical methods (see (3.1)). The discretization (3.1) is a class of one-step methods including the common numerical methods for the linear stochastic oscillator in the literature (see the review article [21]). Based on the Gärtner–Ellis theorem, we first study the LDPs of the mean position and the mean velocity for both the exact solution of the linear stochastic oscillator and its numerical approximations. Then, by giving the conditions which make numerical methods have at least first order convergence in the mean-square sense, we prove that stochastic symplectic methods in the form of (3.1) asymptotically preserve these two LDPs. However, it is shown that neither of the two LDPs is preserved asymptotically by the considered nonsymplectic methods based on the tail estimation of Gaussian random variables. To the best of our knowledge, this is the first result about using LDP to show the superiority of stochastic symplectic methods compared with nonsymplectic ones.

The paper is organized as follows. In section 2, we give some basic concepts about the LDP. For the linear stochastic oscillator, we introduce the mean position $A_T = \frac{1}{T}\int_0^T X_t\,\mathrm{d}t$ and the mean velocity $B_T = \frac{X_T}{T}$ for each $T > 0$ with $\{X_t\}_{t\geq 0}$ being the exact solution. Also, the LDPs for both $\{A_T\}_{T>0}$ and $\{B_T\}_{T>0}$ are established based on the Gärtner–Ellis theorem. Section 3 studies the LDP for the discrete mean position $\{A_N\}_{N\geq 1}$ (see (3.2)) of the general numerical methods for the fixed sufficiently small step-size. In section 4, we derive the pointwise convergence of the modified rate functions of $\{A_N\}_{N\geq 1}$ as step-size tends to zero and show that symplectic methods asymptotically preserve the LDP for $\{A_T\}_{T>0}$. In section 5, by following the ideas of dealing with $\{A_N\}_{N\geq 1}$, we investigate the LDP for the discrete mean velocity $\{B_N\}_{N\geq 1}$ (see (5.1)) and show that symplectic methods asymptotically preserve the LDP for $\{B_T\}_{T>0}$. In section 6, we verify our theoretical results by discussing some concrete numerical methods and construct some methods preserving exactly the LDPs for $\{A_T\}_{T>0}$ or $\{B_T\}_{T>0}$. These imply the superiority of symplectic methods in preserving the LDPs for $\{A_T\}_{T>0}$ and $\{B_T\}_{T>0}$ of the linear stochastic oscillator. We perform numerical experiments to verify our theoretical results in section 7. Finally, in section 8, we give our conclusions and propose several open problems for future study.

2. LDPs for $\{A_T\}_{T>0}$ and $\{B_T\}_{T>0}$.

In this section, we aim to prove that both the mean position $\{A_T\}_{T>0}$ and mean velocity $\{B_T\}_{T>0}$ of the exact solution of our considered stochastic oscillator satisfy the LDPs. Before showing the LDPs of $\{A_T\}_{T>0}$ and $\{B_T\}_{T>0}$, we introduce some preliminaries upon the theory of large deviations, which can be found in [11, 16].

DEFINITION 2.1. *$I : E \to [0,\infty]$ is called a rate function if it is lower semicontinuous, where E is a Polish space. If all level sets $I^{-1}([-\infty, a])$, $a \in [0,\infty)$, are compact, then I is called a good rate function.*

DEFINITION 2.2. *Let I be a rate function and $(\mu_\epsilon)_{\epsilon>0}$ be a family of probability measures on E. We say that $(\mu_\epsilon)_{\epsilon>0}$ satisfies an LDP with the rate function I if*

$$\text{(LDP1)} \qquad \liminf_{\epsilon\to 0} \epsilon \log(\mu_\epsilon(\mathrm{U})) \geq -\inf I(U) \qquad \textit{for every open } U \subset E,$$

$$\text{(LDP2)} \qquad \limsup_{\epsilon\to 0} \epsilon \log(\mu_\epsilon(\mathrm{C})) \leq -\inf I(C) \qquad \textit{for every closed } C \subset E.$$

Based on Definition 2.2, one can give the definition of LDP for a family of random variables similarly. Namely, let $\{X_\epsilon\}_{\epsilon>0}$ be a family of random variables from $(\Omega, \mathscr{F}, \mathbf{P})$ to $(E, \mathscr{B}(E))$. $\{X_\epsilon\}_{\epsilon>0}$ is said to satisfy an LDP with the rate function I if its distribution $(\mathbf{P} \circ X_\epsilon^{-1})_{\epsilon>0}$ satisfies (LDP1) and (LDP2) in Definition 2.2 (see, e.g., [7, 11]).

The Gärtner–Ellis theorem plays an important role in dealing with the LDPs for a family of random variables. When utilizing this theorem, one needs to examine whether the logarithmic moment generating function is essentially smooth.

Definition 2.3. *A convex function $\Lambda : \mathbb{R}^d \to (-\infty, \infty]$ is essentially smooth if*

(1) *$\mathcal{D}_\Lambda^\circ$ is nonempty, where $\mathcal{D}_\Lambda^\circ$ is the interior of $\mathcal{D}_\Lambda := \{x \in \mathbb{R}^d : \Lambda(x) < \infty\}$;*

(2) *$\Lambda(\cdot)$ is differentiable throughout $\mathcal{D}_\Lambda^\circ$;*

(3) *$\Lambda(\cdot)$ is steep, namely, $\lim_{n\to\infty} |\nabla\Lambda(\lambda_n)| = \infty$ whenever $\{\lambda_n\}$ is a sequence in $\mathcal{D}_\Lambda^\circ$ converging to a boundary point of $\mathcal{D}_\Lambda^\circ$.*

Theorem 2.4 (Gärtner–Ellis). *Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of random vectors taking values in $\mathbb{R}^d$. Assume that for each $\lambda \in \mathbb{R}^d$, the logarithmic moment generating function, defined as the limit $\Lambda(\lambda) \triangleq \lim_{n\to\infty} \frac{1}{n} \log(\mathbf{E} e^{n\langle\lambda, X_n\rangle})$, exists as an extended real number. Further, assume that the origin belongs to $\mathcal{D}_\Lambda^\circ$. If Λ is an essentially smooth and lower semicontinuous function, then the LDP holds for $\{X_n\}_{n\in\mathbb{N}}$ with the good rate function $\Lambda^*(\cdot)$. Here $\Lambda^*(x) = \sup_{\lambda\in\mathbb{R}^d}\{\langle\lambda, x\rangle - \Lambda(\lambda)\}$, $x \in \mathbb{R}^d$, is the Fenchel–Legendre transform of $\Lambda(\cdot)$.*

It is known that the key point of the Gärtner–Ellis theorem is to study the logarithmic moment generating function. Moreover, we would like to mention that the Gärtner–Ellis theorem is valid in the case of continuous parameter family $\{X_\epsilon\}_{\epsilon>0}$ (see the remarks of [11, Theorem 2.3.6]).

The motivation of this paper is to explain the superiority of stochastic symplectic methods, by studying the LDPs of numerical methods for a linear stochastic oscillator $\ddot{X}_t + X_t = \alpha\dot{W}_t$ with $\alpha > 0$, and W_t being a one-dimensional standard Brownian motion defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbf{P})$. The linear stochastic oscillator can be rewritten as a two-dimensional stochastic Hamiltonian system

$$\mathrm{d}\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} X_t \\ Y_t \end{pmatrix}\mathrm{d}t + \alpha\begin{pmatrix} 0 \\ 1 \end{pmatrix}\mathrm{d}W_t, \quad \begin{pmatrix} X_0 \\ Y_0 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \tag{2.1}$$

whose phase flow preserves symplectic structure. Namely, the oriented areas of the phase flow are invariant:

$$dX_t \wedge dY_t = dx_0 \wedge dy_0 \qquad \forall \quad t \geq 0,$$

where the exact solution (X_t, Y_t) of (2.1) (see, e.g., [17, Chapter 8]) is

$$\begin{aligned} X_t &= x_0\cos(t) + y_0\sin(t) + \alpha\int_0^t \sin(t-s)\mathrm{d}W_s, \\ Y_t &= -x_0\sin(t) + y_0\cos(t) + \alpha\int_0^t \cos(t-s)\mathrm{d}W_s. \end{aligned} \tag{2.2}$$

To inherit the symplecticity of this stochastic oscillator, different kinds of symplectic methods have been constructed (see [8, 21] and references therein).

For SDE (2.1), we introduce the so-called mean position,

$$A_T = \frac{1}{T}\int_0^T X_t\,\mathrm{d}t \qquad \forall \quad T > 0, \tag{2.3}$$

and the mean velocity,

$$B_T = \frac{X_T}{T} \qquad \forall \quad T > 0. \tag{2.4}$$

Both A_T and B_T are important observables, and they have many applications in physics. For example, the Ornstein–Uhlenbeck process is often used to describe the velocity of a particle moving in a random environment [19]. In this case, A_T can be interpreted as the mean value of the displacement process $\int_0^T X_t \,\mathrm{d}t$, and B_T as the mean value of velocity X_t on the time interval $[0, T]$ (see also [13]). Next, by means of the Gärtner–Ellis theorem, we show that both the mean position $\{A_T\}_{T>0}$ and mean velocity $\{B_T\}_{T>0}$ of the exact solution satisfy the LDPs.

THEOREM 2.5. $\{A_T\}_{T>0}$ *satisfies an LDP with the good rate function* $I(y) = \frac{y^2}{3\alpha^2}$, *i.e.,*

$$\liminf_{T\to\infty} \frac{1}{T}\log(\mathbf{P}(A_T \in U)) \geq -\inf_{y\in U} I(y) \qquad \textit{for every open } U \subset \mathbb{R},$$

$$\limsup_{T\to\infty} \frac{1}{T}\log(\mathbf{P}(A_T \in C)) \leq -\inf_{y\in C} I(y) \qquad \textit{for every closed } C \subset \mathbb{R}.$$

Proof. It follows from (2.2), (2.3), and the stochastic Fubini theorem that

$$TA_T = \int_0^T X_t \mathrm{d}t = x_0\sin(T) + y_0(1-\cos(T)) + \alpha\int_0^T \Big[1-\cos(T-s)\Big]\mathrm{d}W_s. \tag{2.5}$$

Thus, we have $\mathbf{E}\left[TA_T\right] = x_0\sin(T) + y_0(1-\cos(T))$, and

$$\mathbf{Var}\left[TA_T\right] = \alpha^2\int_0^T \Big[1-\cos(T-s)\Big]^2\mathrm{d}s = \alpha^2\left[\frac{3T}{2} - 2\sin(T) + \frac{\sin(2T)}{4}\right].$$

Hence $\lambda TA_T \sim \mathcal{N}(\lambda\mathbf{E}[TA_T], \lambda^2\mathbf{Var}[TA_T])$ for every $\lambda \in \mathbb{R}$. It follows from the characteristic function of λTA_T that $\mathbf{E}e^{\lambda TA_T} = e^{\lambda\mathbf{E}[TA_T] + \frac{\lambda^2}{2}\mathbf{Var}[TA_T]}$. In this way, we obtain the logarithmic moment generating function $\Lambda(\lambda) = \lim_{T\to\infty}\frac{1}{T}\log\mathbf{E}e^{\lambda TA_T} = \frac{3\alpha^2}{4}\lambda^2$, which means that $\Lambda(\cdot)$ is an essentially smooth, lower semicontinuous function. Moreover, we have that the origin 0 belongs to $\mathcal{D}_\Lambda^\circ = \mathbb{R}$. By Theorem 2.4, we obtain that $\{A_T\}_{T>0}$ satisfies an LDP with the good rate function $I(y) = \Lambda^*(y) = \sup_{\lambda\in\mathbb{R}}\{y\lambda - \Lambda(\lambda)\} = \frac{y^2}{3\alpha^2}$. ∎

Notice that LDP for $\{A_T\}_{T>0}$ is independent of the initial value (x_0, y_0) of the stochastic oscillator (2.1). Theorem 2.5 indicates that, for any initial value (x_0, y_0), the probability that the mean position $\{A_T\}_{T>0}$ hits the interval $[a, a+\mathrm{d}a]$ decays exponentially and formally satisfies $\mathbf{P}\left(A_T \in [a, a+\mathrm{d}a]\right) \approx e^{-TI(a)}\mathrm{d}a = e^{-T\frac{a^2}{3\alpha^2}}\mathrm{d}a$ for sufficiently large T.

Next, we give the result of the LDP for $\{B_T\}_{T>0}$ in the following theorem, whose proof is analogous to that of Theorem 2.5 and hence is omitted.

THEOREM 2.6. $\{B_T\}_{T>0}$ *satisfies an LDP with the good rate function* $J(y) = \frac{y^2}{\alpha^2}$.

The above two theorems give the LDPs of $\{A_T\}_{T>0}$ and $\{B_T\}_{T>0}$. For a numerical approximation $\{(x_n, y_n)\}_{n\geq 0}$ of the linear stochastic oscillator (2.1), two natural questions are these: Do its discrete mean position $A_N = \frac{1}{N}\sum_{n=0}^{N-1} x_n$ and discrete mean velocity $B_N = \frac{x_N}{Nh}$ satisfy similar LDPs as in continuous case? Is the method able to asymptotically preserve the LDPs of $\{A_T\}_{T>0}$ and $\{B_T\}_{T>0}$ in the sense that the modified rate functions converge to the rate functions of the exact solution? The next several sections of this paper are devoted to answering the above questions.

3. LDP for discrete mean position $\{A_N\}_{N\geq 1}$. In this section, we study the LDP for the discrete mean position of general numerical methods. We show that symplectic methods and nonsymplectic ones satisfy different types of LDPs based on Theorem 2.4.

Let $\{(x_n, y_n)\}_{n\geq 0}$ be the discrete approximations at $t_n = nh$ with $x_n \approx X_{t_n}$, $y_n \approx Y_{t_n}$, where $h > 0$ is the given step-size. Following [21], we consider the general numerical methods in form of

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = A \begin{pmatrix} x_n \\ y_n \end{pmatrix} + \alpha b \Delta W_n := \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix} + \alpha \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \Delta W_n \tag{3.1}$$

with $\Delta W_n = W_{t_{n+1}} - W_{t_n}$. In fact, the real matrix A and the real vector b depend on both the method and the constant step-size h. In addition, we require $b_1^2 + b_2^2 \neq 0$, which is natural since an effective numerical method for (2.1) must depend on the Brownian motion. In the previous section, we derive the LDP for the mean position $\{A_T\}_{T>0}$ of the continuous system (2.1). In what follows, we consider the LDP for discrete mean position $\{A_N\}_{N\geq 1}$ of the method (3.1) and study how closely the LDP for $\{A_N\}_{N\geq 1}$ approximates the LDP for $\{A_T\}_{T>0}$. We recall that A_N is defined as

$$A_N = \frac{1}{N} \sum_{n=0}^{N-1} x_n, \qquad N = 1, 2, \ldots. \tag{3.2}$$

Our idea is to use Theorem 2.4 to show the LDP of $\{A_N\}_{N\geq 1}$. Hence, we first derive the logarithmic moment generating function $\Lambda^h(\lambda) := \lim_{N\to\infty} \frac{1}{N} \log \mathbf{E} e^{\lambda N A_N}$ for a fixed appropriate step-size h. This can be done by means of the general formula of $\{x_n\}_{n\geq 1}$. For this end, we give a useful lemma (see Appendix A for its proof).

LEMMA 3.1. *For arbitrary $\theta \in (0, 2\pi)$, $N \in \mathbb{N}^+$, and $a \in \mathbb{R}$, it holds that*

$$\sum_{n=1}^{N} \sin(n\theta) a^n = \frac{a\sin(\theta) - a^{N+1}\sin((N+1)\theta) + a^{N+2}\sin(N\theta)}{1 - 2a\cos(\theta) + a^2}. \tag{3.3}$$

In particular, if $a = 1$, then

$$\sum_{n=1}^{N} \sin(n\theta) = \frac{\cos\left(\frac{\theta}{2}\right) - \cos((N+\frac{1}{2})\theta)}{2\sin\left(\frac{\theta}{2}\right)}. \tag{3.4}$$

In order to give the general formula of $\{x_n\}_{n\geq 1}$, we suppose

$$4\det(A) - (\mathrm{tr}(A))^2 > 0. \tag{A1}$$

Then using Lemma 3.1, we have the following proposition.

PROPOSITION 3.2. *For the method* (3.1) *satisfying the assumption* (A1), *we have that for any $n, N \geq 1$,*

$$\begin{aligned} x_n = {} & \left(a_{11}\hat{\alpha}_{n-1} + \hat{\beta}_{n-1}\right) x_0 + a_{12}\hat{\alpha}_{n-1} y_0 \\ & + \alpha \sum_{j=0}^{n-1} \left[b_1 \hat{\alpha}_{n-1-j} + (a_{12}b_2 - a_{22}b_1)\hat{\alpha}_{n-2-j}\right] \Delta W_j, \end{aligned} \tag{3.5}$$

$$N A_N = \left(1 + a_{11} S_N^{\hat{\alpha}} + S_N^{\hat{\beta}}\right) x_0 + a_{12} S_N^{\hat{\alpha}} y_0 + \alpha \sum_{j=0}^{N-2} c_j \Delta W_j. \tag{3.6}$$

Here, $\hat{\alpha}_k = (\det(A))^{k/2} \frac{\sin((k+1)\theta)}{\sin(\theta)}$ and $\hat{\beta}_k = -(\det(A))^{(k+1)/2} \frac{\sin(k\theta)}{\sin(\theta)}$ for any integer k, *with* $\theta \in (0, \pi)$ *satisfying*

$$\cos(\theta) = \frac{\mathrm{tr(A)}}{2\sqrt{\det(A)}}, \qquad \sin(\theta) = \frac{\sqrt{4\det(A) - (\mathrm{tr(A)})^2}}{2\sqrt{\det(A)}}. \tag{3.7}$$

And

$$S_N^{\hat{\alpha}} = \frac{\sin(\theta) - \left(\sqrt{\det(A)}\right)^{N-1}\sin(N\theta) + \left(\sqrt{\det(A)}\right)^{N}\sin((N-1)\theta)}{\sin(\theta)\left(1 - 2\sqrt{\det(A)}\cos(\theta) + \det(A)\right)}, \tag{3.8}$$

$$S_N^{\hat{\beta}} = -\frac{\det(A)\sin(\theta) - \left(\sqrt{\det(A)}\right)^{N}\sin((N-1)\theta) + \left(\sqrt{\det(A)}\right)^{N+1}\sin((N-2)\theta)}{\sin(\theta)\left(1 - 2\sqrt{\det(A)}\cos(\theta) + \det(A)\right)}, \tag{3.9}$$

$$\begin{aligned} c_j = {} & \frac{b_1}{\sin(\theta)}\sin((N-1-j)\theta)\left(\sqrt{\det(A)}\right)^{N-2-j} + \frac{b_1 + a_{12}b_2 - a_{22}b_1}{\sin(\theta)} \\ & \times \frac{\sin(\theta) - \left(\sqrt{\det(A)}\right)^{N-2-j}\sin((N-1-j)\theta) + \left(\sqrt{\det(A)}\right)^{N-1-j}\sin((N-2-j))\theta)}{1 - 2\sqrt{\det(A)}\cos(\theta) + \det(A)}. \end{aligned} \tag{3.10}$$

Proof. Denote $M_n = \binom{x_{n+1}}{x_n}$ for $n \geq 1$. It follows from the recurrence (3.1) that $M_n = DM_{n-1} + r_n$, $n \geq 1$ with

$$D = \begin{pmatrix} \mathrm{tr}(A) & -\det(A) \\ 1 & 0 \end{pmatrix}, \quad r_n = \begin{pmatrix} \alpha\left(b_1\Delta W_n + (a_{12}b_2 - a_{22}b_1)\Delta W_{n-1}\right) \\ 0 \end{pmatrix},$$

where $\mathrm{tr}(A)$ and $\det(A)$ denote the trace and the determinant of A, respectively. In this way, we have $M_n = D^n M_0 + \sum_{j=1}^n D^{n-j} r_j$, $n \geq 1$. Under the assumption (A1), matrix D has two complex-valued eigenvalues

$$\lambda_\pm = \frac{\mathrm{tr}(A)}{2} \pm \boldsymbol{i}\frac{\sqrt{4\det(A) - (\mathrm{tr}(A))^2}}{2} = \sqrt{\det(A)}e^{\pm \boldsymbol{i}\theta}, \qquad \boldsymbol{i}^2 = -1,$$

where θ satisfies (3.7).

It follows from the expression of M_n (one can refer to [21]) that

$$x_{n+1} = \hat{\alpha}_n x_1 + \hat{\beta}_n x_0 + \alpha\sum_{j=1}^n \hat{\alpha}_{n-j}\left[b_1\Delta W_j + (a_{12}b_2 - a_{22}b_1)\Delta W_{j-1}\right], \qquad n \geq 0.$$

Since $x_1 = a_{11}x_0 + a_{12}y_0 + \alpha b_1\Delta W_0$, $\hat{\alpha}_{-1} = 0$, and $\hat{\alpha}_0 = 1$, for $n \geq 1$,

$$\begin{aligned} x_n &= \left(a_{11}\hat{\alpha}_{n-1} + \hat{\beta}_{n-1}\right)x_0 + a_{12}\hat{\alpha}_{n-1}y_0 + \alpha b_1\hat{\alpha}_{n-1}\Delta W_0 \\ &\quad + \alpha\sum_{j=1}^{n-1} b_1\hat{\alpha}_{n-1-j}\Delta W_j + \alpha\sum_{j=0}^{n-2}(a_{12}b_2 - a_{22}b_1)\hat{\alpha}_{n-2-j}\Delta W_j \\ &= \left(a_{11}\hat{\alpha}_{n-1} + \hat{\beta}_{n-1}\right)x_0 + a_{12}\hat{\alpha}_{n-1}y_0 \\ &\quad + \alpha\sum_{j=0}^{n-1}\left[b_1\hat{\alpha}_{n-1-j} + (a_{12}b_2 - a_{22}b_1)\hat{\alpha}_{n-2-j}\right]\Delta W_j. \end{aligned}$$

By (3.2) and (3.5), we have

$$NA_N = x_0 + \sum_{n=1}^{N-1} x_n = \left(1 + a_{11}S_N^{\hat{\alpha}} + S_N^{\hat{\beta}}\right) x_0 + a_{12}S_N^{\hat{\alpha}} y_0 + \alpha \sum_{j=0}^{N-2} c_j \Delta W_j, \tag{3.11}$$

where $S_N^{\hat{\alpha}} = \sum_{n=0}^{N-2} \hat{\alpha}_n$, $S_N^{\hat{\beta}} = \sum_{n=0}^{N-2} \hat{\beta}_n$ and

$$\begin{aligned} c_j &:= \sum_{n=j+1}^{N-1} \Big[b_1 \hat{\alpha}_{n-1-j} + (a_{12}b_2 - a_{22}b_1)\hat{\alpha}_{n-2-j}\Big] \\ &= b_1 \hat{\alpha}_{N-2-j} + (b_1 + a_{12}b_2 - a_{22}b_1) S_{N-1-j}^{\hat{\alpha}}. \end{aligned} \tag{3.12}$$

Finally, by means of (3.3), we obtain the explicit expressions of $S_N^{\hat{\alpha}}$, $S_N^{\hat{\beta}}$, and c_j as given in (3.8), (3.9), and (3.10), respectively. This completes the proof. ∎

Next, we study the LDP of $\{A_N\}_{N\geq 1}$ for symplectic methods and nonsymplectic ones, respectively. It is known that the method (3.1) preserves the symplectic structure, i.e., $dx_{n+1} \wedge dy_{n+1} = dx_n \wedge dy_n$, if and only if $\det(A) = 1$. (In fact, this condition is equivalent to the fact that method (3.1) preserves the phase volume.) In addition, for nonsymplectic methods, we exclude the case $\det(A) > 1$, where the logarithmic moment generating function Λ^h does not exist. This is to say, we need to deal with the case $\det(A) = 1$ and the case $\det(A) < 1$ separately.

3.1. LDP of $\{A_N\}_{N\geq 1}$ for symplectic methods. In this part, we derive the LDP for $\{A_N\}_{N\geq 1}$ of the method (3.1) in the case of preserving the symplecticity. By (3.6), NA_N is Gaussian. Hence, $\Lambda^h(\lambda) = \lim_{N\to\infty} \frac{1}{N} \log \mathbf{E} e^{\lambda N A_N} = \lim_{N\to\infty} \frac{1}{N}[\lambda \mathbf{E}(NA_N) + \frac{1}{2}\lambda^2 \mathbf{Var}(NA_N)]$. In order to get the expression of Λ^h, it suffices to give the estimates of $\mathbf{E}(NA_N)$ and $\mathbf{Var}(NA_N)$ with respect to N.

Hereafter we use the notation $K(a_1, \dots, a_m)$ to denote some constant dependent on the parameters $a_1, \dots, a_m$ but independent of N, which may vary from one line to another. We assume that

$$\det(A) = 1. \tag{A2}$$

Under (A2), we have $\hat{\alpha}_n = \frac{\sin((n+1)\theta)}{\sin(\theta)}$, $\hat{\beta}_n = -\frac{\sin(n\theta)}{\sin(\theta)}$. Then by (3.4) and (3.12), we obtain

$$S_N^{\hat{\alpha}} = \frac{\cos\left(\frac{\theta}{2}\right) - \cos((N-\frac{1}{2})\theta)}{2\sin(\theta)\sin\left(\frac{\theta}{2}\right)}, \quad S_N^{\hat{\beta}} = -\frac{\cos\left(\frac{\theta}{2}\right) - \cos((N-\frac{3}{2})\theta)}{2\sin(\theta)\sin\left(\frac{\theta}{2}\right)}, \tag{3.13}$$

$$c_j = \frac{(b_1 + a_{12}b_2 - a_{22}b_1)\cos\left(\frac{\theta}{2}\right) - b_1\cos((N-\frac{1}{2}-j)\theta) - (a_{12}b_2 - a_{22}b_1)\cos((N-\frac{3}{2}-j)\theta)}{2\sin(\theta)\sin\left(\frac{\theta}{2}\right)}. \tag{3.14}$$

By (3.13), it holds that $|S_N^{\hat{\alpha}}| + |S_N^{\hat{\beta}}| \leq K(\theta)$ for each $N \geq 2$. Further, it follows from (3.6) and (3.14) that

$$|\mathbf{E}[NA_N]| = \left|\left(1 + a_{11}S_N^{\hat{\alpha}} + S_N^{\hat{\beta}}\right) x_0 + a_{12}S_N^{\hat{\alpha}} y_0\right| \leq K(x_0, y_0, \theta), \tag{3.15}$$

$$\mathbf{Var}[NA_N] = \alpha^2 h \sum_{j=0}^{N-2} c_j^2 = \frac{\alpha^2 h}{4\sin^2(\theta)\sin^2\left(\frac{\theta}{2}\right)} \sum_{j=0}^{N-2} \tilde{c}_j^2 \tag{3.16}$$

with

$$\tilde{c}_j^2 = (b_1+a_{12}b_2-a_{22}b_1)^2\cos^2\left(\frac{\theta}{2}\right)+\frac{1}{2}b_1^2+\frac{1}{2}(a_{12}b_2-a_{22}b_1)^2+b_1(a_{12}b_2-a_{22}b_1)\cos(\theta)+R_j,$$

where

$$\begin{aligned} R_j = {} & \frac{b_1^2}{2}\cos((2N-1-2j)\theta) + \frac{(a_{12}b_2-a_{22}b_1)^2}{2}\cos((2N-3-2j)\theta) \\ & - 2b_1(b_1+a_{12}b_2-a_{22}b_1)\cos\left(\frac{\theta}{2}\right)\cos\left(\frac{(2N-1-2j)\theta}{2}\right) \\ & - 2(b_1+a_{12}b_2-a_{22}b_1)(a_{12}b_2-a_{22}b_1)\cos\left(\frac{\theta}{2}\right)\cos\left(\frac{(2N-3-2j)\theta}{2}\right) \\ & + b_1(a_{12}b_2-a_{22}b_1)\cos((2N-2-2j)\theta). \end{aligned}$$

We claim $|\sum_{j=0}^{N-2} R_j| \le K(\theta)$. In detail, by $\sum_{n=1}^{N}\cos((2n+1)\theta) = \frac{\sin((2N+2)\theta)-\sin(2\theta)}{2\sin(\theta)}$, we have

$$\left|\sum_{j=0}^{N-2}\cos((2N-1-2j)\theta)\right| = \left|\sum_{n=1}^{N-1}\cos((2n+1)\theta)\right| = \left|\frac{\sin(2N\theta)-\sin(2\theta)}{2\sin(\theta)}\right| \le K(\theta).$$

Analogously, we obtain

$$\begin{aligned} & \left|\sum_{j=0}^{N-2}\cos((2N-3-2j)\theta)\right| + \left|\sum_{j=0}^{N-2}\cos(\frac{(2N-1-2j)\theta}{2})\right| \\ & + \left|\sum_{j=0}^{N-2}\cos(\frac{(2N-3-2j)\theta}{2})\right| + \left|\sum_{j=0}^{N-2}\cos((2N-2-2j)\theta)\right| \le K(\theta), \end{aligned} \tag{3.17}$$

which proves the above claim.

Based on (3.15), (3.16), (3.1), and $|\sum_{j=0}^{N-2} R_j| \le K(\theta)$, we have

$$\begin{aligned} \Lambda^h(\lambda) & := \lim_{N\to\infty}\frac{1}{N}\log \mathbf{E}e^{\lambda N A_N} \\ & = \frac{\alpha^2 h\lambda^2}{8\sin^2(\theta)\sin^2\left(\frac{\theta}{2}\right)}\bigg[(b_1+a_{12}b_2-a_{22}b_1)^2\cos^2\left(\frac{\theta}{2}\right)+\frac{1}{2}b_1^2 \\ & \qquad + \frac{1}{2}(a_{12}b_2-a_{22}b_1)^2 + b_1(a_{12}b_2-a_{22}b_1)\cos(\theta)\bigg]. \end{aligned} \tag{3.18}$$

As a result of (3.7) with $\det(A)=1$, it holds that

$$\begin{aligned} & \cos(\theta) = \frac{\mathrm{tr}(A)}{2}, \quad \sin(\theta) = \frac{\sqrt{4-(\mathrm{tr}(A))^2}}{2}, \\ & \sin^2\left(\frac{\theta}{2}\right) = \frac{1-\cos(\theta)}{2} = \frac{2-\mathrm{tr}(A)}{4}, \quad \cos^2\left(\frac{\theta}{2}\right) = \frac{1+\cos(\theta)}{2} = \frac{2+\mathrm{tr}(A)}{4}. \end{aligned} \tag{3.19}$$

Substituting (3.19) into (3.18) yields that

$$\Lambda^h(\lambda) = \frac{\alpha^2 h \lambda^2}{2(2+\mathrm{tr}(A))(2-\mathrm{tr}(A))^2}\Big[(b_1 + a_{12}b_2 - a_{22}b_1)^2(4+\mathrm{tr}(A)) - 2b_1(a_{12}b_2 - a_{22}b_1)(2-\mathrm{tr}(A))\Big]. \tag{3.20}$$

In order to show that Λ^h is essentially smooth. We need to use the following lemma, whose proof is given in Appendix B.

LEMMA 3.3. *Under assumptions* (A1) *and* (A2), *we have*
(1) $b_1^2 + (a_{12}b_2 - a_{22}b_1)^2 \neq 0$;
(2) $(b_1 + a_{12}b_2 - a_{22}b_1)^2(4+\mathrm{tr(A)}) - 2\mathrm{b_1(a_{12}b_2 - a_{22}b_1)(2-tr(A))} > 0.$

Lemma 3.3(2) means that Λ^h is essentially smooth. It follows from Theorem 2.4 that $\{A_N\}_{N\geq 1}$ satisfies an LDP with the good rate function

$$\begin{aligned} I^h(y) &= \sup_{\lambda\in\mathbb{R}}\{y\lambda - \Lambda^h(\lambda)\} \\ &= \frac{(2+\mathrm{tr}(A))(2-\mathrm{tr}(A))^2 y^2}{2\alpha^2 h\Big[(b_1 + a_{12}b_2 - a_{22}b_1)^2(4+\mathrm{tr}(A)) - 2b_1(a_{12}b_2 - a_{22}b_1)(2-\mathrm{tr}(A))\Big]}. \end{aligned} \tag{3.21}$$

Finally, we acquire the following theorem.

THEOREM 3.4. *If the numerical method* (3.1) *for approximating the SDE* (2.1) *satisfies assumptions* (A1) *and* (A2), *then its mean position* $\{A_N\}_{N\geq 1}$ *satisfies an LDP with the good rate function given by* (3.21).

Remark 3.5. Theorem 3.4 indicates that to make the LDP hold for $\{A_N\}_{N\geq 1}$, the step-size h needs to be restricted such that assumptions (A1) and (A2) hold. Moreover, the rate function $I^h(y)$ does not depend on the initial value (x_0, y_0). That is to say, for appropriate step-size h and arbitrary initial value, $\{A_N\}_{N\geq 1}$ formally satisfies $\mathbf{P}(A_N \in [a, a+\mathrm{d}a]) \approx e^{-NI^h(a)}\mathrm{d}a$ for sufficiently large N.

3.2. LDP of $\{A_N\}_{N\geq 1}$ for nonsymplectic methods. In this part, we show the LDP for $\{A_N\}_{N\geq 1}$ of method (3.1) under the following assumption:

$$0 < \det(A) < 1. \tag{A3}$$

Under the assumption (A3), one immediately concludes from (3.8), (3.9), and the boundedness of $\sin(\cdot)$ that $|S_N^{\hat{\alpha}}| + |S_N^{\hat{\beta}}| \leq K(\theta, A)$ for all $N \geq 2$, which gives

$$\Big|\mathbf{E}[NA_N]\Big| \leq K(x_0, y_0, \theta, A). \tag{3.22}$$

It follows from (3.6) and (3.10) that

$$\mathbf{Var}(NA_N) = \alpha^2 h \sum_{j=0}^{N-2} c_j^2, \tag{3.23}$$

where

$$c_j^2 = \left(\frac{b_1 + a_{12}b_2 - a_{22}b_1}{1 - 2\sqrt{\det(A)}\cos(\theta) + \det(A)}\right)^2 + \tilde{R}_j \tag{3.24}$$

with

$$
\begin{aligned}
\tilde{R}_j = {} & \frac{b_1^2 \sin^2((N-1-j)\theta)(\det(A))^{N-2-j}}{\sin^2(\theta)} + \frac{(b_1 + a_{12}b_2 - a_{22}b_1)^2}{\sin^2(\theta)\left(1 - 2\sqrt{\det(A)}\cos(\theta) + \det(A)\right)^2} \\
& \times \Bigg[(\det(A))^{N-2-j}\sin^2((N-1-j)\theta) + (\det(A))^{N-1-j}\sin^2((N-2-j)\theta) \\
& \quad - 2\sin(\theta)\left(\sqrt{\det(A)}\right)^{N-2-j}\sin((N-1-j)\theta) \\
& \quad + 2\sin(\theta)\left(\sqrt{\det(A)}\right)^{N-1-j}\sin((N+2-j)\theta) \\
& \quad - 2\left(\sqrt{\det(A)}\right)^{2N-3-2j}\sin((N-1-j)\theta)\sin((N-2-j)\theta)\Bigg] \\
& + \frac{2b_1(b_1 + a_{12}b_2 - a_{22}b_1)}{\sin^2(\theta)\left(1 - 2\sqrt{\det(A)}\cos(\theta) + \det(A)\right)} \\
& \Bigg[\sqrt{\det(A)}^{N-2-j}\sin(\theta)\sin((N-1-j)\theta) - (\det(A))^{N-2-j}\sin^2((N-1-j)\theta)) \\
& \quad + \left(\sqrt{\det(A)}\right)^{2N-3-2j}\sin((N-1-j)\theta)\sin((N-2-j)\theta)\Bigg].
\end{aligned}
$$

Further, we have

$$
\begin{aligned}
\sum_{j=0}^{N-2}\left[(\det(A))^{N-2-j} + (\det(A))^{N-1-j}\right] &\leq 2\sum_{j=0}^{N}(\det(A))^j, \\
\sum_{j=0}^{N-2}\left[\left(\sqrt{\det(A)}\right)^{N-2-j} + \left(\sqrt{\det(A)}\right)^{N-1-j}\right] &\leq 2\sum_{j=0}^{N}\left(\sqrt{\det(A)}\right)^j, \\
\sum_{j=0}^{N-2}\left(\sqrt{\det(A)}\right)^{2N-3-2j} = \sum_{j=1}^{N-1}\left(\sqrt{\det(A)}\right)^{2j-1} &\leq \frac{1}{\sqrt{\det(A)}}\sum_{j=0}^{N}(\det(A))^j.
\end{aligned}
$$

It follows from the boundedness of $\sin(\cdot)$ and $\det(A) < 1$ that

$$
\left|\sum_{j=0}^{N-2}\tilde{R}_j\right| \leq K(\theta, A)\sum_{j=0}^{N}\left[\left(\sqrt{\det(A)}\right)^j + (\det(A))^j\right] \leq \widetilde{K}(\theta, A), \tag{3.25}
$$

where $\widetilde{K}(\theta, A)$ is a constant dependent on θ and A but independent of N. Combining (3.22), (3.23), (3.24), and (3.25) leads to

$$
\begin{aligned}
\widetilde{\Lambda}^h(\lambda) &= \lim_{N\to\infty}\frac{1}{N}\log \mathbf{E}e^{\lambda N A_N} \\
&= \frac{\alpha^2 h\lambda^2}{2}\lim_{N\to\infty}\frac{1}{N}\left[\left(\frac{b_1 + a_{12}b_2 - a_{22}b_1}{1 - 2\sqrt{\det(A)}\cos(\theta) + \det(A)}\right)^2 (N-1) + \sum_{j=0}^{N-2}\tilde{R}_j\right] \\
&= \frac{\alpha^2 h\lambda^2}{2}\left(\frac{b_1 + a_{12}b_2 - a_{22}b_1}{1 - 2\sqrt{\det(A)}\cos(\theta) + \det(A)}\right)^2.
\end{aligned}
$$

If we assume that

$$b_1 + a_{12}b_2 - a_{22}b_1 \neq 0, \tag{A4}$$

then it follows from Theorem 2.4 that $\{A_N\}_{N\geq 1}$ satisfies an LDP with the good rate function $\widetilde{I}^h(y) = \frac{y^2}{2\alpha^2 h}(\frac{1-2\sqrt{\det(A)}\cos(\theta)+\det(A)}{b_1+a_{12}b_2-a_{22}b_1})^2 = \frac{y^2}{2\alpha^2 h}(\frac{1-\mathrm{tr}(A)+\det(A)}{b_1+a_{12}b_2-a_{22}b_1})^2$, where we have used (3.7) in the second equality. Finally, we obtain the following theorem.

THEOREM 3.6. *If the numerical method* (3.1) *for approximating the SDE* (2.1) *satisfies assumptions* (A1), (A3), *and* (A4), *then its mean position* $\{A_N\}_{N\geq 1}$ *satisfies an LDP with the good rate function* $\widetilde{I}^h(y) = \frac{y^2}{2\alpha^2 h}(\frac{1-\mathrm{tr}(A)+\det(A)}{b_1+a_{12}b_2-a_{22}b_1})^2$.

4. Asymptotical preservation for the LDP of $\{A_T\}_{T>0}$.

In section 3, we acquire the LDP for the discrete mean position $\{A_N\}_{N\geq 1}$ when the method (3.1) is symplectic or nonsymplectic separately, for given appropriate step-size. In this section, we study their asymptotical preservation for the LDP of $\{A_T\}_{T>0}$ as step-size tends to 0 (see Definition 1.1). By Definition 1.1, we obtain the modified rate functions of the rate functions appearing in Theorems 3.4 and 3.6, respectively, as follows:

$$I^h_{mod}(y) = \frac{(2+\mathrm{tr}(A))(2-\mathrm{tr}(A))^2 y^2}{2\alpha^2 h^2\left[(b_1+a_{12}b_2-a_{22}b_1)^2(4+\mathrm{tr}(A)) - 2b_1(a_{12}b_2-a_{22}b_1)(2-\mathrm{tr}(A))\right]}, \tag{4.1}$$

$$\widetilde{I}^h_{mod}(y) = \frac{y^2}{2\alpha^2 h^2}\left(\frac{1-\mathrm{tr}(A)+\det(A)}{b_1+a_{12}b_2-a_{22}b_1}\right)^2. \tag{4.2}$$

It would fail to get the asymptotical convergence for $I^h_{mod}(y)$ and $\widetilde{I}^h_{mod}(y)$ only by means of assumptions (A1)–(A4) in two aspects: one is that both A and b are some functions of step-size h, which are unknown unless a specific method is applied; the other is that for some A and b, the numerical approximation may not be convergent to the original system. A solution to this problem is studying the convergence on finite interval of numerical methods. In what follows, we consider the mean-square convergence of the method (3.1).

For the sake of simplicity, we first give some notation. Let $R = \mathcal{O}(h^p)$ stand for $|R| \leq Ch^p$ for all sufficiently small step-size h, where C is independent of h and may vary from one line to another. $f(h) \sim h^p$ means that $f(h)$ and h^p are equivalent infinitesimal. Furthermore, $\|\cdot\|_2$ denotes the 2-norm of a vector or matrix and $\|\cdot\|_F$ denotes the Frobenius norm of a matrix.

Since (2.1) is driven by the additive noise, the mean-square convergence order of general numerical methods which are known for the moment to approximate this system is no less than 1. Hence, in what follows, we restrict (3.1) to the numerical method with at least first order convergence in the mean-square sense. To give the conditions about the mean-square convergence of the method (3.1), we introduce the Euler–Maruyama method of form (3.1) with $A^{EM} = \begin{pmatrix} 1 & h \\ -h & 1 \end{pmatrix}$, $b^{EM} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Based on the fundamental convergence theorem, we acquire the sufficient conditions which make numerical method (3.1) have at least first order convergence in the mean-square sense.

THEOREM 4.1. *If the numerical method* (3.1) *satisfies*

$$\left\|A - A^{EM}\right\|_F = \mathcal{O}(h^2) \qquad \textit{and} \qquad \left\|b - b^{EM}\right\|_2 = \mathcal{O}(h), \tag{4.3}$$

then its convergence order is at least 1 *in the mean-square sense on any finite interval* $[0, T_0]$*, i.e.,* $\sup_{n\geq 0,\ nh\leq T_0}[\mathbf{E}((x_n - X(t_n))^2 + (y_n - Y(t_n))^2)]^{1/2} \leq K(T_0)h$*.*

We put the proof of this theorem into Appendix C. In fact, it is verified that (4.3) is equivalent to

$$\text{(B)}\qquad |a_{11} - 1| + |a_{22} - 1| + |a_{12} - h| + |a_{21} + h| = \mathcal{O}(h^2), \text{ and } |b_1| + |b_2 - 1| = \mathcal{O}(h).$$

Using the assumption (B), we have the following lemma (its proof is given in Appendix D), which is used to study whether method (3.1) asymptotically preserves the LDPs for $\{A_T\}_{T>0}$ or $\{B_T\}_{T>0}$ of the exact solution.

Lemma 4.2. *Under the assumption* (B)*, the following properties hold:*
(1) $\mathrm{tr}(A) \to 2$ *as* $h \to 0$*;*
(2) $(1 - \mathrm{tr}(A) + \det(A)) \sim h^2$*;*
(3) $(b_1 + a_{12}b_2 - a_{22}b_1) \sim h$*.*

By Lemma 4.2, we obtain the convergence of the modified rate functions in (4.1) and (4.2).

Case 1: Let (A1), (A2), and (B) hold. Noting $\det(A) = 1$ in this case, Lemma 4.2(2) yields $(2 - \mathrm{tr}(A)) \sim h^2$. Hence,

$$\lim_{h\to 0} \frac{b_1(a_{12}b_2 - a_{22}b_1)\,(2 - \mathrm{tr}(A))}{h^2} = 0. \tag{4.4}$$

It follows from Lemma 4.2, (4.1), and (4.4) that

$$\begin{aligned}\lim_{h\to 0} I^h_{mod}(y) &= \frac{y^2}{2\alpha^2}\frac{\lim_{h\to 0}(2 + \mathrm{tr}(A))}{\lim_{h\to 0}(4 + \mathrm{tr}(A))(b_1 + a_{12}b_2 - a_{22}b_1)^2/h^2 - 2\lim_{h\to 0} b_1(a_{12}b_2 - a_{22}b_1)(2 - \mathrm{tr}(A))/h^2} \\ &= \frac{y^2}{3\alpha^2}.\end{aligned} \tag{4.5}$$

Case 2: Let (A1), (A3), (A4), and (B) hold. According to (4.2) and Lemma 4.2, we have $\lim_{h\to 0} \widetilde{I}^h_{mod} = \frac{y^2}{2\alpha^2} \lim_{h\to 0} \frac{(h^2)^2}{h^2\cdot h^2} = \frac{y^2}{2\alpha^2}$. Therefore, by Definition 1.1, we get the following two theorems.

Theorem 4.3. *For the numerical method* (3.1) *approximating the stochastic oscillator* (2.1)*, if assumptions* (A1) *and* (A2) *hold, then we have the following:*

(1) *The method* (3.1) *is symplectic.*

(2) *The discrete mean position* $\{A_N\}_{N\geq 1}$ *of method* (3.1) *satisfies an LDP with the good rate function*

$$I^h(y) = \frac{(2 + \mathrm{tr}(A))(2 - \mathrm{tr}(A))^2 y^2}{2\alpha^2 h\Big[(b_1 + a_{12}b_2 - a_{22}b_1)^2(4 + \mathrm{tr}(A)) - 2b_1(a_{12}b_2 - a_{22}b_1)(2 - \mathrm{tr}(A))\Big]}. \tag{4.6}$$

(3) *Moreover, if the assumption* (B) *holds, then method* (3.1) *asymptotically preserves the LDP of* $\{A_T\}_{T>0}$*, i.e., the modified rate function* $I^h_{mod}(y) = I^h(y)/h$ *satisfies*

$$\lim_{h\to 0} I^h_{mod}(y) = I(y) \qquad \forall \quad y \in \mathbb{R},$$

where $I(\cdot)$ *is the rate function of LDP for* $\{A_T\}_{T>0}$*.*

THEOREM 4.4. *For the numerical method* (3.1) *approximating the stochastic oscillator* (2.1), *if assumptions* (A1), (A3), *and* (A4) *hold, then we have the following:*

(1) *The method* (3.1) *is nonsymplectic.*

(2) *The discrete mean position* $\{A_N\}_{N\geq 1}$ *of method* (3.1) *satisfies an LDP with the good rate function* $\widetilde{I}^h(y) = \frac{y^2}{2\alpha^2 h}(\frac{1-\mathrm{tr}(A)+\det(A)}{b_1+a_{12}b_2-a_{22}b_1})^2$.

(3) *Moreover, if the assumption* (B) *holds, then method* (3.1) *does not asymptotically preserve the LDP of* $\{A_T\}_{T>0}$, *i.e., for* $y \neq 0$, $\lim_{h\to 0}\widetilde{I}^h_{mod}(y) \neq I(y)$, *where* $\widetilde{I}^h_{mod}(y) = \widetilde{I}^h(y)/h$, *and* $I(\cdot)$ *is the rate function of LDP for* $\{A_T\}_{T>0}$.

Remark 4.5. Theorems 4.3 and 4.4 indicate that under appropriate conditions, the symplectic methods asymptotically preserve the LDP for the mean position $\{A_T\}_{T>0}$ of original system (2.1), while the nonsymplectic methods do not. This implies that, in comparison with nonsymplectic methods, symplectic methods have long-time stability in the aspect of preserving the LDP for the mean position.

5. LDP for discrete mean velocity $\{B_N\}_{N\geq 1}$.

In section 2, we obtain the LDP for the mean velocity $\{B_T\}_{T>0}$ of the original system (2.1). In this section, following the ideas of dealing with the discrete mean position, we investigate the LDP for the discrete mean velocity. That is to say, we first prove the LDP of the discrete mean velocity based on the Gärtner–Ellis theorem for a fixed appropriate step-size h. Then, we derive the limits of the modified rate functions by means of the assumption (B).

We consider the numerical approximation of $B_T = \frac{X_T}{T}$ at $t_N = Nh$. Noting that x_N is used to approximate X_{t_N} in terms of the numerical method (3.1), we define discrete mean velocity as

$$B_N = \frac{x_N}{Nh}, \qquad N = 1, 2, \ldots. \tag{5.1}$$

In what follows, we study the LDP for $\{B_N\}_{N\geq 1}$ of method (3.1) and its asymptotical preservation for LDP of $\{B_T\}_{T>0}$. Similar to the arguments on $\{A_N\}_{N\geq 1}$, we introduce the modified rate function to characterize how the LDP for $\{B_N\}_{N\geq 1}$ approximates the LDP for $\{B_T\}_{T>0}$.

We still assume that (A1) holds. In this case, the equality (3.5) holds. Then

$$\begin{aligned} x_N = {} & \left(a_{11}\hat{\alpha}_{N-1} + \hat{\beta}_{N-1}\right)x_0 + a_{12}\hat{\alpha}_{N-1}y_0 \\ & + \alpha\sum_{n=0}^{N-1}\left[b_1\hat{\alpha}_{N-1-n} + (a_{12}b_2 - a_{22}b_1)\hat{\alpha}_{N-2-n}\right]\Delta W_n \end{aligned} \tag{5.2}$$

with

$$\hat{\alpha}_n = (\det(A))^{n/2}\frac{\sin((n+1)\theta)}{\sin(\theta)}, \qquad \hat{\beta}_n = -\left(\det(A)\right)^{\frac{n+1}{2}}\frac{\sin(n\theta)}{\sin(\theta)}.$$

According to (5.2), x_N is a Gaussian random variable whose expectation is

$$\begin{aligned} \mathbf{E}(x_N) = {} & \left(a_{11}\left(\det(A)\right)^{\frac{N-1}{2}}\frac{\sin(N\theta)}{\sin(\theta)} - \left(\det(A)\right)^{\frac{N}{2}}\frac{\sin((N-1)\theta)}{\sin(\theta)}\right)x_0 \\ & + a_{12}\left(\det(A)\right)^{\frac{N-1}{2}}\frac{\sin(N\theta)}{\sin(\theta)}y_0. \end{aligned}$$

If $0 < \det(A) \leq 1$, then $|\mathbf{E}(x_N)| \leq K(\theta)$, which leads to

$$\lim_{N\to\infty}\frac{\mathbf{E}(x_N)}{N} = 0. \tag{5.3}$$

From (5.2) and the fact $\hat{\alpha}_{-1} = 0$, we get

$$\begin{aligned}\mathbf{Var}(x_N) &= \alpha^2 h \sum_{n=0}^{N-1} \Big[b_1 \hat{\alpha}_{N-1-n} + (a_{12}b_2 - a_{22}b_1)\hat{\alpha}_{N-2-n}\Big]^2 \\ &= \alpha^2 h \Bigg[\Big(b_1^2 + (a_{12}b_2 - a_{22}b_1)^2\Big) \sum_{n=0}^{N-2} \hat{\alpha}_n^2 + b_1^2 \hat{\alpha}_{N-1} \\ &\qquad +2b_1(a_{12}b_2 - a_{22}b_1) \sum_{n=1}^{N-1} \hat{\alpha}_n \hat{\alpha}_{n-1}\Bigg]. \end{aligned} \tag{5.4}$$

Further, we have

$$\sum_{n=0}^{N-2} \hat{\alpha}_n^2 = \sum_{n=0}^{N-2} \frac{(\det(A))^n \sin^2((n+1)\theta)}{\sin^2(\theta)}, \tag{5.5}$$

$$2\sum_{n=1}^{N-1} \hat{\alpha}_n \hat{\alpha}_{n-1} = \frac{1}{\sin^2(\theta)} \sum_{n=1}^{N-1} (\det(A))^{\frac{2n-1}{2}} (\cos(\theta) - \cos((2n+1)\theta)). \tag{5.6}$$

As is analogous to the treatment of $\{A_N\}_{N\ge1}$, we deal with symplectic methods ($\det(A) = 1$) and nonsymplectic ones ($0 < \det(A) < 1$), respectively.

5.1. LDP of $\{B_N\}_{N\ge1}$ for symplectic methods. In this part, we study the LDP for $\{B_N\}_{N\ge1}$ of symplectic methods, so we assume that (A2) holds. Based on $\det(A) = 1$, (5.5), and (5.6), we have

$$\sum_{n=0}^{N-2} \hat{\alpha}_n^2 = \frac{1}{\sin^2(\theta)} \sum_{n=1}^{N-1} \sin^2(n\theta) = \frac{1}{\sin^2(\theta)} \left(\frac{N-1}{2} - \frac{\sin((2N-1)\theta) - \sin(\theta)}{4\sin(\theta)}\right) \tag{5.7}$$

and

$$2\sum_{n=1}^{N-1} \hat{\alpha}_n \hat{\alpha}_{n-1} = \frac{1}{\sin^2(\theta)} \left[(N-1)\cos(\theta) - \frac{\sin(2N\theta) - \sin(2\theta)}{2\sin(\theta)}\right]. \tag{5.8}$$

Substituting (5.7) and (5.8) into (5.4) yields

$$\begin{aligned}\mathbf{Var}(x_N) = \alpha^2 h \Bigg[&\frac{b_1^2 + (a_{12}b_2 - a_{22}b_1)^2 + 2b_1(a_{12}b_2 - a_{22}b_1)\cos(\theta)}{2\sin^2(\theta)}(N-1) \\ &- \frac{\big[b_1^2 + (a_{12}b_2 - a_{22}b_1)^2\big]\left[\sin((2N-1)\theta) - \sin(\theta)\right]}{4\sin^3(\theta)} + \frac{b_1^2 \sin^2(N\theta)}{\sin^2(\theta)} \\ &- \frac{b_1(a_{12}b_2 - a_{22}b_1)(\sin(2N\theta) - \sin(\theta))}{2\sin^3(\theta)}\Bigg]. \end{aligned} \tag{5.9}$$

Using (5.3), (5.9), and (3.7) with $\det(A) = 1$, we have

$$\begin{aligned}\Lambda^h(\lambda) &= \lim_{N\to\infty} \frac{1}{N} \log \mathbf{E} e^{\lambda N B_N} \\ &= \frac{\alpha^2\lambda^2 \left[(b_1 + a_{12}b_2 - a_{22}b_1)^2 - b_1(a_{12}b_2 - a_{22}b_1)(2 - \mathrm{tr}(A))\right]}{(4 - (\mathrm{tr}(A))^2)\, h}. \end{aligned} \tag{5.10}$$

Before proving that Λ^h is essentially smooth, we give the following lemma (see its proof in Appendix E).

LEMMA 5.1. *Under assumptions* (A1) *and* (A2), *it holds that* $(b_1+a_{12}b_2-a_{22}b_1)^2-b_1(a_{12}b_2-a_{22}b_1)(2-\mathrm{tr}(\mathrm{A}))>0$.

Lemma 5.1 shows that $\Lambda^h(\cdot)$ is essentially smooth and lower semicontinuous. Then, using Theorem 2.4, we conclude that $\{B_N\}_{N\geq 1}$ satisfies an LDP with the good rate function

$$J^h(y)=\frac{h\left[4-(\mathrm{tr}(A))^2\right]y^2}{4\alpha^2\Big[(b_1+a_{12}b_2-a_{22}b_1)^2-b_1(a_{12}b_2-a_{22}b_1)(2-\mathrm{tr}(A))\Big]}. \tag{5.11}$$

By Definition 1.1, the modified rate function is

$$J^h_{mod}(y)=\frac{\big(4-(\mathrm{tr}(A))^2\big)y^2}{4\alpha^2\Big[(b_1+a_{12}b_2-a_{22}b_1)^2-b_1(a_{12}b_2-a_{22}b_1)(2-\mathrm{tr}(A))\Big]}. \tag{5.12}$$

In what follows, we study the asymptotical convergence of $J^h_{mod}(\cdot)$ as step-size h tends to 0 based on the mean-square convergence condition. To this end, let the assumption (B) hold. Then it follows from Lemma 4.2 that $(2-\mathrm{tr}(A))\sim h^2$, $(b_1+a_{12}b_2-a_{22}b_1)\sim h$. In addition, (B) implies that $b_1(a_{12}b_2-a_{22}b_1)\to 0$ as $h\to 0$. In this way, we have

$$\lim_{h\to 0}J^h_{mod}(y)=\frac{2+\lim_{h\to 0}\mathrm{tr}(A)}{4\alpha^2\left[\lim_{h\to 0}\frac{(b_1+a_{12}b_2-a_{22}b_1)^2}{2-\mathrm{tr}(A)}-\lim_{h\to 0}b_1(a_{12}b_2-a_{22}b_1)\right]}y^2=\frac{y^2}{\alpha^2}.$$

According to the above results, we write them into the following theorem.

THEOREM 5.2. *For the numerical method* (3.1) *approximating the stochastic oscillator* (2.1), *if assumptions* (A1) *and* (A2) *hold, then we have the following:*

(1) *The method* (3.1) *is symplectic.*

(2) *The discrete mean velocity* $\{B_N\}_{N\geq 1}$ *of method* (3.1) *satisfies an LDP with the good rate function*

$$J^h(y)=\frac{h\left[4-(\mathrm{tr}(\mathrm{A}))^2\right]y^2}{4\alpha^2\Big[(b_1+a_{12}b_2-a_{22}b_1)^2-b_1(a_{12}b_2-a_{22}b_1)(2-\mathrm{tr}(\mathrm{A}))\Big]}.$$

(3) *Moreover, if the assumption* (B) *holds, then method* (3.1) *asymptotically preserves the LDP of* $\{B_T\}_{T>0}$, *i.e., the modified rate function* $J^h_{mod}(y)=J^h(y)/h$ *satisfies*

$$\lim_{h\to 0}J^h_{mod}(y)=J(y)\qquad\forall\quad y\in\mathbb{R},$$

where $J(\cdot)$ *is the rate function of the LDP for* $\{B_T\}_{T>0}$.

5.2. LDP of $\{B_N\}_{N\geq 1}$ for nonsymplectic methods. In this part, we consider the discrete mean velocity $\{B_N\}_{N\geq 1}$ of general nonsymplectic methods. We study whether the LDP holds for $\{B_N\}_{N\geq 1}$. Let assumptions (A1) and (A3) hold. Then, (5.5) and (5.6) satisfy, respectively,

$$\sum_{n=0}^{N-2}\hat{\alpha}_n^2\leq K(\theta)\sum_{n=0}^{N-2}(\det(A))^n\leq K_1(\theta),$$

$$\left|2\sum_{n=1}^{N-1}\hat{\alpha}_n\hat{\alpha}_{n-1}\right|\leq K(\theta)\sum_{n=1}^{N-1}(\det(A))^{\frac{2n-1}{2}}\leq K_2(\theta),$$

where $K_1(\theta)$ and $K_2(\theta)$ are two constants dependent on θ but independent of N. Additionally, it holds that $|\hat{\alpha}_{N-1}| = |\frac{(\det(A))^{N-1}\sin^2(N\theta)}{\sin^2(\theta)}| \leq K(\theta)$. Thus, (5.4) satisfies

$$|\mathbf{Var}(x_N)| \leq \alpha^2 h K(\theta). \tag{5.13}$$

It follows from (5.3) and (5.13) that the logarithmic moment generating function is

$$\widetilde{\Lambda}^h(\lambda) = \lim_{N\to\infty} \frac{1}{N} \log \mathbf{E} e^{\lambda N B_N} = \lim_{N\to\infty} \frac{1}{N}\left[\frac{\lambda}{h}\mathbf{E}(x_N) + \frac{\lambda^2}{2h^2}\mathbf{Var}(x_N)\right] = 0. \tag{5.14}$$

We note that $\widetilde{\Lambda}^h(\cdot)$ is not essentially smooth, for which Theorem 2.4 is not valid. In our case, we can directly prove that the LDP holds for $\{B_N\}_{N\geq 1}$ of nonsymplectic methods by the definition of LDP. We claim that $\{B_N\}_{N\geq 1}$ of nonsymplectic methods satisfies the LDP with the good rate function:

$$\tilde{J}^h(y) = \begin{cases} 0, & y = 0,\\ +\infty, & y \neq 0. \end{cases} \tag{5.15}$$

We divide the proof of this claim into three steps.

Step 1: *We show the limit behaviors of* $\mathbf{P}(B_N \geq x_0)$ *and* $\mathbf{P}(B_N \leq x_0)$ *for nonsymplectic methods.*

We need to use the following fact: if $X \sim \mathcal{N}(\mu, \sigma^2)$, then it follows from [16, Lemma 22.2] that, for any $x > \mu$,

$$\mathbf{P}\left(X \geq x\right) = \mathbf{P}\left(\frac{X-\mu}{\sigma} \geq \frac{x-\mu}{\sigma}\right) \leq \frac{1}{\sqrt{2\pi}}\frac{\sigma}{x-\mu}e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{5.16}$$

In addition, for any $x < \mu$,

$$\mathbf{P}\left(X \leq x\right) = \mathbf{P}\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = \mathbf{P}\left(\frac{X-\mu}{\sigma} \geq -\frac{x-\mu}{\sigma}\right) \leq \frac{1}{\sqrt{2\pi}}\frac{\sigma}{\mu-x}e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{5.17}$$

Since $B_N = \frac{x_N}{Nh}$, we have $B_N \sim \mathcal{N}(\frac{\mathbf{E}(x_N)}{Nh}, \frac{\mathbf{Var}(x_N)}{N^2h^2})$ with $|\mathbf{E}(x_N)| \leq K(\theta)$ and $|\mathbf{Var}(x_N)| \leq K(\theta)$. Noting that $\lim_{N\to\infty}\mathbf{E}(B_N) = 0$, one has that for the given $x_0 > 0$, there exists some N_0 such that $\mathbf{E}(B_N) < x_0$ for every $N > N_0$. Accordingly, it follows from (5.16) that

$$\mathbf{P}\left(B_N \geq x_0\right) \leq \frac{1}{\sqrt{2\pi}}\frac{\sqrt{\mathbf{Var}(x_N)}}{Nhx_0 - \mathbf{E}(x_N)}\exp\left\{-\frac{\left(Nhx_0 - \mathbf{E}(x_N)\right)^2}{2\mathbf{Var}(x_N)}\right\} \qquad \forall \quad N > N_0.$$

In this way, for every $x_0 > 0$,

$$\lim_{N\to\infty}\frac{1}{N}\log\left[\mathbf{P}\left(B_N \geq x_0\right)\right] = -\infty. \tag{5.18}$$

Analogously, using (5.17), one has that for the given $x_0 < 0$,

$$\lim_{N\to\infty}\frac{1}{N}\log\left[\mathbf{P}\left(B_N \leq x_0\right)\right] = -\infty. \tag{5.19}$$

Step 2: *We prove the upper bound LDP (LDP2): For every closed* $C \subset \mathbb{R}$,

$$\limsup_{N\to\infty}\frac{1}{N}\log\mathbf{P}(B_N \in C) \leq -\inf \tilde{J}^h(C). \tag{5.20}$$

If $0 \in C$, then it follows from (5.15) that $\inf \tilde{J}^h(C) = 0$. Since $\mathbf{P}(B_N \in C) \leq 1$, (5.20) naturally holds.

If $0 \notin C$, define $x_+ = \inf(C \bigcap (0, +\infty))$ and $x_- = \sup(C \bigcap (-\infty, 0))$. Then, $\mathbf{P}(B_N \in C) \leq \mathbf{P}(B_N \geq x_+) + \mathbf{P}(B_N \leq x_-)$. In order to prove (5.20), we need to use the following lemma (see [16, Lemma 23.9]).

Lemma 5.3. *Let $N \in \mathbb{N}$ and a_ϵ^i, $i = 1, \dots, N$, $\epsilon > 0$, be nonnegative numbers. Then* $\limsup_{\epsilon \to 0} \epsilon \log \sum_{i=1}^N a_\epsilon^i = \max_{i=1,\dots,N} \limsup_{\epsilon\to 0} \epsilon \log(a_\epsilon^i)$.

Using (5.18), (5.19), and Lemma 5.3 yields

$$\begin{aligned}
&\limsup_{N\to\infty} \frac{1}{N} \log \mathbf{P}(B_N \in C) \\
&\qquad \leq \max\left\{ \limsup_{N\to\infty} \frac{1}{N} \log \mathbf{P}(B_N \geq x_+),\ \limsup_{N\to\infty} \frac{1}{N} \log \mathbf{P}(B_N \leq x_-) \right\} = -\infty.
\end{aligned}$$

Noting that $0 \notin C$, one obtains $\inf \tilde{J}^h(C) = +\infty$. Thus, (5.20) also holds for this case.

Step 3: *We prove the lower bound LDP (LDP1): For every open $U \subset \mathbb{R}$,*

$$\liminf_{N\to\infty} \frac{1}{N} \log \mathbf{P}(B_N \in U) \geq -\inf \tilde{J}^h(U). \tag{5.21}$$

If $0 \notin U$, then $\inf \tilde{J}^h(U) = +\infty$. Since $\mathbf{P}(B_N \in C) \geq 0$, (5.21) naturally holds.

If $0 \in U$, then there exists some $\delta > 0$ such that $(-\delta, \delta) \subset U$. Accordingly,

$$\liminf_{N\to\infty} \frac{1}{N} \log \mathbf{P}(B_N \in U) \geq \liminf_{N\to\infty} \frac{1}{N} \log \mathbf{P}(|B_N| < \delta). \tag{5.22}$$

It follows from (5.18) that for arbitrary given $M \in (-\infty, 0)$, there exists some N_1 such that for every $N > N_1$, $\frac{1}{N} \log [\mathbf{P}(B_N \geq \delta)] < M$. Thus,

$$\mathbf{P}(B_N \geq \delta) \leq e^{NM} \qquad \forall \quad N > N_1,$$

which leads to $\lim_{N\to\infty} \mathbf{P}(B_N \geq \delta) = 0$. Similarly, utilizing (5.19) gives $\lim_{N\to\infty} \mathbf{P}(B_N \leq -\delta) = 0$. Hence, $\lim_{N\to\infty} \mathbf{P}(|B_N| < \delta) = 1$, which implies

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbf{P}(|B_N| < \delta) = 0. \tag{5.23}$$

Combining (5.22) and (5.23), we have $\liminf_{N\to\infty} \frac{1}{N} \log \mathbf{P}(B_N \in U) \geq 0$. Further, since $0 \in U$, $\inf \tilde{J}^h(U) = 0$. Hence, we prove (5.21).

Combining the above discussion, we deduce that $\{B_N\}_{N\geq 1}$ of nonsymplectic methods satisfies the LDP with the good rate function $\tilde{J}^h$ given by (5.15) and the modified rate function $\tilde{J}^h_{mod} = \tilde{J}^h/h = \tilde{J}^h$. Finally, we get the following theorem.

Theorem 5.4. *For the numerical method* (3.1) *approximating the stochastic oscillator* (2.1), *if assumptions* (A1) *and* (A3) *hold, then we have the following:*

(1) *The method* (3.1) *is nonsymplectic.*

(2) *The discrete mean velocity $\{B_N\}_{N\geq 1}$ of method* (3.1) *satisfies an LDP with the good rate function* $\tilde{J}^h(y) = \begin{cases} 0, & y=0, \\ +\infty, & y\neq 0. \end{cases}$

(3) *Method* (3.1) *does not asymptotically preserve the LDP of $\{B_T\}_{T>0}$, i.e., for $y \neq 0$, $\lim_{h\to 0} \tilde{J}^h_{mod}(y) \neq J(y)$, where $\tilde{J}^h_{mod}(y) = \tilde{J}^h(y)/h$, and $J(y) = \frac{y^2}{\alpha^2}$ is the rate function of LDP for $\{B_T\}_{T>0}$.*

6. Concrete numerical methods. In this section, we show and compare the LDPs of some concrete numerical methods to verify the theoretical results obtained in previous sections. For symplectic methods, we consider the symplectic β-method, the exponential method, INT method, and OPT method. For nonsymplectic ones, we examine the θ-method, PC (PEM-MR) method, and PC (EM-BEM) method. All of the methods can be found in [21], except the symplectic β-method (see, e.g., [18, equation (2.7)]). Furthermore, we construct some symplectic methods which preserve the LDP for $\{A_T\}_{T>0}$ or $\{B_T\}_{T>0}$ exactly.

6.1. Symplectic methods.

• Symplectic β-method ($\beta \in [0,1]$):

$$A^\beta = \frac{1}{1+\beta(1-\beta)h^2}\begin{pmatrix} 1-(1-\beta)^2h^2 & h \\ -h & 1-\beta^2h^2 \end{pmatrix},$$
$$b^\beta = \frac{1}{1+\beta(1-\beta)h^2}\begin{pmatrix} (1-\beta)h \\ 1 \end{pmatrix}.$$

The straightforward calculation leads to

$$\det(A^\beta) = 1, \qquad \operatorname{tr}(A^\beta) = \frac{2-(2\beta^2-2\beta+1)h^2}{1+\beta(1-\beta)h^2}, \tag{6.1}$$

$$a_{12}b_2 - a_{22}b_1 = \frac{\beta h}{1+\beta(1-\beta)h^2}, \qquad b_1 + a_{12}b_2 - a_{22}b_1 = \frac{h}{1+\beta(1-\beta)h^2}. \tag{6.2}$$

It can be verified that the assumption (B) holds, and if $h \in (0,2)$, then for every $\beta \in [0,1]$, assumptions (A1) and (A2) hold. Substituting (6.1) and (6.2) into (4.6), we have $I^h(y) = \frac{hy^2}{3\alpha^2}[\frac{3}{2} - \frac{3}{6-(2\beta-1)^2h^2}]$, which is the good rate function of LDP for $\{A_N\}_{N\geq 1}$ of the symplectic β-method by Thoerem 4.3. Furthermore, we get the modified rate function $I^h_{mod}(y) = I^h(y)/h = \frac{y^2}{3\alpha^2}[\frac{3}{2} - \frac{3}{6-(2\beta-1)^2h^2}]$.

Further, we have that $\lim_{h\to 0} I^h_{mod}(y) = I(y) = \frac{y^2}{3\alpha^2}$ for every $y \in \mathbb{R}$, which is consistent with the third conclusion of Theorem 4.3. Moreover, for every $h > 0$, the modified rate function of the mean position for the midpoint method with $\beta = \frac{1}{2}$ is the same as that for the exact solution. These indicate that the midpoint method exactly preserves the LDP for $\{A_T\}_{T>0}$. In the case of $\beta \neq \frac{1}{2}$, $I^h_{mod}(y) < I(y)$ provided $y \neq 0$. That is, as the time T and t_N tend to infinity simultaneously, the exponential decay speed of $\mathbf{P}(A_N \in [a, a+\mathrm{d}a])$ is slower than that of $\mathbf{P}(A_T \in [a, a+\mathrm{d}a])$ provided $a \neq 0$.

On the other hand, if $h \in (0,2)$ and $\beta \in (0,1)$, assumptions (A1) and (A2) hold. By Theorem 5.2, $\{B_N\}_{N\geq 1}$ of the symplectic β-method satisfies an LDP with the good rate function $J^h(y) = \frac{h[4-(2\beta-1)^2h^2][1+\beta(1-\beta)h^2]y^2}{4\alpha^2}$. This means that the modified rate function $J^h_{mod}(\cdot)$ satisfies $\lim_{h\to 0} J^h_{mod}(y) = \frac{y^2}{\alpha^2} = J(y)$, which verifies the third conclusion of Theorem 5.2.

• Exponential method (EX): $A^{EX} = \begin{pmatrix} \cos(h) & \sin(h) \\ -\sin(h) & \cos(h) \end{pmatrix}$, $\quad b^{EX} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

For this method, it holds that

$$\det(A^{EX}) = 1, \ \operatorname{tr}(A^{EX}) = 2\cos(h),$$
$$a_{12}b_2 - a_{22}b_1 = \sin(h), \ b_1 + a_{12}b_2 - a_{22}b_1 = \sin(h).$$

If $h \in (0,\pi)$, then assumptions (A1) and (A2) hold. Then, we obtain that $\{A_N\}_{N\geq 1}$

satisfies an LDP with the modified rate function $I^h_{mod}(y) = \frac{2y^2}{\alpha^2}\frac{1-\cos(h)}{h^2(2+\cos(h))}$. Hence, we have $\lim_{h\to 0} I^h_{mod}(y) = \frac{y^2}{3\alpha^2} = I(y)$. One can show that $I^h_{mod}(y) > I(y)$ provided that $h \in (0, \pi/6)$ and $y \neq 0$.

According to the discussions above, if $h \in (0, \pi/6)$, then the mean position $\{A_N\}_{N\geq 1}$ of the exponential method satisfies an LDP, which asymptotically preserves the LDP for $\{A_T\}_{T>0}$. In addition, as the time T and t_N tend to infinity simultaneously, the exponential decay speed of $\mathbf{P}(A_N \in [a, a+\mathrm{d}a])$ is faster than that of $\mathbf{P}(A_T \in [a, a + \mathrm{d}a])$ provided that $a \neq 0$.

Analogously, we have that assumptions (A1) and (A2) hold for $h \in (0, \pi)$. Hence, for $h \in (0, \pi)$, $\{B_N\}_{N\geq 1}$ of the exponential method satisfies an LDP with the modified rate function $J^h_{mod}(y) = \frac{y^2}{\alpha^2} = J(y)$. In this way, the exponential method exactly preserves the LDP for $\{B_T\}_{T>0}$.

- Integral method (INT): $A^{INT} = \begin{pmatrix} \cos(h) & \sin(h) \\ -\sin(h) & \cos(h) \end{pmatrix}$, $b^{INT} = \begin{pmatrix} \sin(h) \\ \cos(h) \end{pmatrix}$.

For this method, $\det(A^{INT}) = 1$, $\operatorname{tr}(A^{INT}) = 2\cos(h)$, $a_{12}b_2 - a_{22}b_1 = 0$ and $b_1 + a_{12}b_2 - a_{22}b_1 = \sin(h)$. It is shown that its modified rate functions of $\{A_N\}_{N\geq 1}$ and $\{B_N\}_{N\geq 1}$ are $I^h_{mod}(y) = \frac{2y^2}{\alpha^2}\frac{1-\cos(h)}{h^2(2+\cos(h))}$, and $J^h_{mod}(y) = \frac{y^2}{\alpha^2} = J(y)$, respectively. This case is exactly the same as that of the exponential method.

- Optimal method (OPT): $A^{OPT} = \begin{pmatrix} \cos(h) & \sin(h) \\ -\sin(h) & \cos(h) \end{pmatrix}$, $b^{OPT} = \frac{1}{h}\begin{pmatrix} 2\sin^2(\frac{h}{2}) \\ \sin(h) \end{pmatrix}$.

Based on the above two formulas, one has

$$\det(A^{OPT}) = 1, \qquad \operatorname{tr}(A^{OPT}) = 2\cos(h), \qquad a_{12}b_2 - a_{22}b_1 = b_1 = \frac{1-\cos(h)}{h}.$$

If $h \in (0, \pi)$, then assumptions (A1) and (A2) hold such that $\{A_N\}_{N\geq 1}$ of the optimal method satisfies an LDP with the modified rate function $I^h_{mod}(y) = \frac{y^2}{3\alpha^2} = I(y)$. Thus, we conclude that the LDP for mean position $\{A_N\}_{N\geq 1}$ of the optimal method exactly preserves the LDP for $\{A_T\}_{T>0}$.

Assumptions (A1) and (A2) hold provided that $h \in (0, \pi)$. Thus, for $h \in (0, \pi)$, $\{B_N\}_{N\geq 1}$ of the optimal method satisfies an LDP with the modified rate function $J^h_{mod}(y) = \frac{h^2y^2}{2(1-\cos(h))\alpha^2}$. Further, we have that $\lim_{h\to 0} J^h_{mod}(y) = \frac{y^2}{\alpha^2} = J(y)$ and $J^h_{mod}(y) > J(y)$. Hence, the optimal method asymptotically preserves the LDP for $\{B_T\}_{T>0}$. When the time T and t_N tend to infinity simultaneously, the exponential decay speed of $\mathbf{P}(B_N \in [a, a+\mathrm{d}a])$ is faster than that of $\mathbf{P}(B_T \in [a, a+\mathrm{d}a])$ provided $a \neq 0$.

6.2. Nonsymplectic methods.

- Stochastic θ-method ($\theta \in [0, 1/2) \cup (1/2, 1]$):

$$A^\theta = \frac{1}{1+\theta^2h^2}\begin{pmatrix} 1-(1-\theta)\theta h^2 & h \\ -h & 1-(1-\theta)\theta h^2 \end{pmatrix}, \qquad b^\theta = \frac{1}{1+\theta^2h^2}\begin{pmatrix} \theta h \\ 1 \end{pmatrix}.$$

For this method, we have

$$\det(A^\theta) = \frac{1+(1-\theta)^2h^2}{1+\theta^2h^2},$$
$$1 - \operatorname{tr}(A^\theta) + \det(A^\theta) = \frac{h^2}{1+\theta^2h^2}, \; b_1 + a_{12}b_2 - a_{22}b_1 = \frac{h}{1+\theta^2h^2}.$$

Notice that $0 < \det(A^\theta) < 1$ is equivalent to $\theta \in (1/2, 1]$. One can show that, with $\theta \in (1/2, 1]$, (A1), (A3), and (A4) hold for every $h > 0$. Hence, for every $\theta \in (1/2, 1]$ and $h > 0$, the mean position $\{A_N\}_{N\geq 1}$ satisfies an LDP with the modified rate function $\widetilde{I}^h_{mod}(y) = \frac{y^2}{2\alpha^2}$, which verifies the third conclusion of Theorem 4.4.

• PC (PEM-MR): $A^1 = \begin{pmatrix} 1-h^2/2 & h(1-h^2/2) \\ -h & 1-h^2/2 \end{pmatrix}$, $\qquad b^1 = \begin{pmatrix} h/2 \\ 1 \end{pmatrix}$.

One has that $1 - \mathrm{tr}(A^1) + \det(A^1) = h^2 - \frac{h^4}{4}$ and $b_1 + a_{12}b_2 - a_{22}b_1 = h - \frac{h^3}{4}$. We obtain that (A1), (A3), and (A4) hold, provided $h \in (0, \sqrt{2})$. Thus, by Theorem 4.4, $\{A_N\}_{N\geq 1}$ of this method satisfies an LDP with the modified rate function $\widetilde{I}^h_{mod}(y) = \frac{y^2}{2\alpha^2}$.

• PC (EM-BEM): $A^2 = \begin{pmatrix} 1-h^2 & h \\ -h & 1-h^2 \end{pmatrix}$, $b^2 = \begin{pmatrix} h \\ 1 \end{pmatrix}$.

We have that $1-\mathrm{tr}(A^2)+\det(A^2) = h^2+h^4$ and $b_1+a_{12}b_2-a_{22}b_1 = h+h^3$. In this case, (A1), (A3), and (A4) hold, provided $h \in (0, 1)$. Thus, by Theorem 4.4, $\{A_N\}_{N\geq 1}$ of this method satisfies an LDP with the modified rate function $\widetilde{I}^h_{mod}(y) = \frac{y^2}{2\alpha^2}$.

We observe that all methods shown in sections 6.1 and 6.2 satisfy the assumption (B). When the step-size h is sufficiently small, the symplectic methods in section 6.1 satisfy assumptions (A1) and (A2), and the nonsymplectic methods in section 6.2 satisfy assumptions (A1), (A3), and (A4). By studying these methods, we verify the theoretical results in Theorems 4.3, 4.4, and 5.2. It is shown that symplectic methods are superior to nonsymplectic methods in terms of preservation of the LDP for both $\{A_T\}_{T>0}$ and $\{B_T\}_{T>0}$.

6.3. Construction for methods exactly preserving the LDP for $\{A_T\}_{T>0}$ or $\{B_T\}_{T>0}$. In this part, we construct several symplectic methods exactly preserving the LDP for $\{A_T\}_{T>0}$ (resp., $\{B_T\}_{T>0}$) based on Theorem 4.3 (resp., Theorem 5.2).

• Methods exactly preserving the LDP for $\{A_T\}_{T>0}$.

Motivated by the assumption (B), we consider the method (3.1) with

$$A = \begin{pmatrix} 1 + c_{11}h^2 & h + c_{12}h^2 \\ -h + c_{21}h^2 & 1 + c_{22}h^2 \end{pmatrix}, \qquad b = \begin{pmatrix} D_1 h \\ 1 + D_2 h \end{pmatrix} \tag{6.3}$$

with real constants c_{ij} and D_i, $i, j = 1, 2$, independent of h. In order to make the condition $\det(A) = 1$ hold, we have

$$(1 + c_{11}h^2)(1 + c_{22}h^2) = 1 + (h + c_{12}h^2)(-h + c_{21}h^2) \quad \forall \quad h > 0.$$

Comparing the coefficients, we obtain

$$c_{11} + c_{22} = -1, \quad c_{11}c_{22} = c_{12}c_{21}, \quad c_{12} = c_{21}.$$

Letting $c_{12} = c_{21} = \sigma$, then c_{11} and c_{22} are the roots of equation $x^2 + x + \sigma^2 = 0$. To assure that c_{11} and c_{22} are real numbers, we assume $\sigma \in [-1/2, 1/2]$. Solving the equation $x^2 + x + \sigma^2 = 0$ yields $c_{11} = \frac{-1-\sqrt{1-4\sigma^2}}{2}$, $c_{22} = \frac{-1+\sqrt{1-4\sigma^2}}{2}$ or $c_{11} = \frac{-1+\sqrt{1-4\sigma^2}}{2}$, $c_{22} = \frac{-1-\sqrt{1-4\sigma^2}}{2}$, where the case $c_{11} = c_{22} = -1/2$, $\sigma = \pm 1/2$ is included in the above two cases. In order to acquire the methods exactly preserving the LDP for $\{A_T\}_{T>0}$, a necessary condition is that the modified rate function (4.1) satisfies

$$I^h_{mod}(y) = \frac{(2+\mathrm{tr}(A))(2-\mathrm{tr}(A))^2y^2}{2\alpha^2h^2\Big[(b_1+a_{12}b_2-a_{22}b_1)^2(4+\mathrm{tr}(A))-2b_1(a_{12}b_2-a_{22}b_1)(2-\mathrm{tr}(A))\Big]} = \frac{y^2}{3\alpha^2}. \tag{6.4}$$

According to (6.3), it is known that

$$\mathrm{tr}(A)=2-h^2,\quad a_{12}b_2-a_{22}b_1 = h\left[(1-D_1)+(D_2+\sigma)h+(D_2\sigma-D_1c_{22})h^2\right].$$

Substituting the above equation into (6.4), we have

$$\begin{aligned} 6-\frac{3h^2}{2} &= \Big[1+(D_2+\sigma)h+(D_2\sigma-D_1c_{22})h^2\Big]^2(6-h^2) \\ &\quad -2D_1h^2\Big[1-D_1+(D_2+\sigma)h+(D_2\sigma-D_1c_{22})h^2\Big]. \end{aligned} \tag{6.5}$$

By comparing the coefficients of h^6 and h^4 in (6.5) and some direct computation, we finally obtain

$$D_1=\frac{1}{2},\quad \sigma=0,\pm\frac{1}{2},\quad c_{22}=\frac{-1+\sqrt{1-4\sigma^2}}{2},\quad c_{11}=\frac{-1-\sqrt{1-4\sigma^2}}{2},\quad D_2=-\sigma.$$

Finally, we acquire three numerical methods, which exactly preserve the LDP for $\{A_T\}_{T>0}$, with the matrix A and the vector b being respectively given by

$$A^{[1]}=\begin{pmatrix} 1-h^2 & h \\ -h & 1\end{pmatrix},\qquad b^{[1]}=\begin{pmatrix} h/2 \\ 1\end{pmatrix}; \tag{6.6}$$

$$A^{[2]}=\begin{pmatrix} 1-h^2/2 & h+h^2/2 \\ -h+h^2/2 & 1-h^2/2\end{pmatrix},\qquad b^{[2]}=\begin{pmatrix} h/2 \\ 1-h/2\end{pmatrix}; \tag{6.7}$$

$$A^{[3]}=\begin{pmatrix} 1-h^2/2 & h-h^2/2 \\ -h-h^2/2 & 1-h^2/2\end{pmatrix},\qquad b^{[3]}=\begin{pmatrix} h/2 \\ 1+h/2\end{pmatrix}. \tag{6.8}$$

Moreover, if $h\in(0,2)$, methods based on (6.6), (6.7), and (6.8) satisfy assumptions (A1) and (A2) and have the same modified rate function $I^h_{mod}(y)=\frac{y^2}{3\alpha^2}=I(y)$.

• Methods exactly preserving the LDP for $\{B_T\}_{T>0}$:

We still consider the method with coefficients satisfying (6.3). By the straightforward computation, we get the following methods exactly preserving the LDP for $\{B_T\}_{T>0}$, whose coefficients are

$$A=\begin{pmatrix} 1-\frac{1+\sqrt{1-4\sigma^2}}{2}h^2 & h+\sigma h^2 \\ -h+\sigma h^2 & 1-\frac{1-\sqrt{1-4\sigma^2}}{2}h^2\end{pmatrix},\qquad b=\begin{pmatrix} h/2 \\ 1-\sigma h\end{pmatrix},$$

with $\sigma=0,\pm\frac{1}{2}$, or

$$A=\begin{pmatrix} 1-\frac{1-\sqrt{1-4\sigma^2}}{2}h^2 & h+\sigma h^2 \\ -h+\sigma h^2 & 1-\frac{1+\sqrt{1-4\sigma^2}}{2}h^2\end{pmatrix},\qquad b=\begin{pmatrix} -h/2 \\ 1-\sigma h\end{pmatrix},$$

with $\sigma = 0, \pm\frac{1}{2}$. Finally, besides methods based on (6.6), (6.7), and (6.8), we obtain three more methods exactly preserving the LDP for $\{B_T\}_{T>0}$ with coefficients given by

$$A^{[4]} = \begin{pmatrix} 1 & h \\ -h & 1-h^2 \end{pmatrix}, \qquad b^{[4]} = \begin{pmatrix} -h/2 \\ 1 \end{pmatrix}; \tag{6.9}$$

$$A^{[5]} = \begin{pmatrix} 1-h^2/2 & h+h^2/2 \\ -h+h^2/2 & 1-h^2/2 \end{pmatrix}, \qquad b^{[5]} = \begin{pmatrix} -h/2 \\ 1-h/2 \end{pmatrix}; \tag{6.10}$$

$$A^{[6]} = \begin{pmatrix} 1-h^2/2 & h-h^2/2 \\ -h-h^2/2 & 1-h^2/2 \end{pmatrix}, \qquad b^{[6]} = \begin{pmatrix} -h/2 \\ 1+h/2 \end{pmatrix}. \tag{6.11}$$

In fact, it is verified that methods based on (6.6)–(6.11) satisfy assumptions (A1) and (A2) for $h \in (0,2)$ and have the same modified rate function $J^h_{mod}(y) = \frac{y^2}{\alpha^2} = J(y)$.

Remark 6.1. Note that three symplectic methods constructed based on (6.6), (6.7), and (6.8) preserve exactly the LDP for $\{A_T\}_{T>0}$ and $\{B_T\}_{T>0}$ at the same time.

7. Numerical experiments. In this section, we perform numerical experiments to verify the theoretical results. We apply the algorithm in [20] to numerically simulate the large deviations rate functions of $\{A_T\}_{T>0}$ and $\{B_T\}_{T>0}$, where the key is to simulate the logarithmic moment generating function based on the Monte–Carlo method.

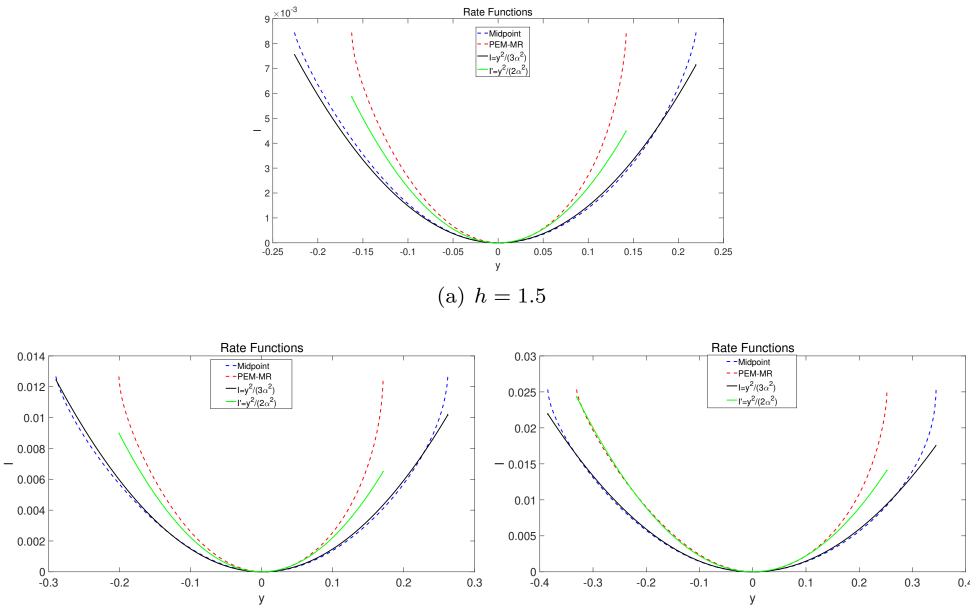
In detail, for a given numerical method $\{x_n, y_n\}_{n\geq 0}$ approximating (2.1), we first obtain M samplings of $\{x_n\}_{n=0}^{N_0-1}$ for a given N_0, which immediately generate M samplings $A^{(i)}_{N_0}$, $i = 1, 2, \ldots, M$ (recall $A_{N_0} = \frac{1}{N_0}\sum_{n=0}^{N_0-1} x_n$). Then we take $G_{M,N_0}(\lambda) = \frac{1}{M}\sum_{i=1}^{M} e^{\lambda N_0 A^{(i)}_{N_0}}$ as the approximation of $\mathbf{E}e^{\lambda N_0 A_{N_0}}$. Further, for sufficiently large N_0, $\Lambda^h_{M,N_0}(\lambda) = \frac{1}{N_0}\log G_{M,N_0}(\lambda)$ is used to approximate $\Lambda^h(\lambda) = \lim_{N\to\infty}\frac{1}{N}\log \mathbf{E}e^{\lambda N A_N}$. Finally, noting that $(G_{M,N_0})'(\lambda) = \frac{1}{M}\sum_{i=1}^{M} e^{N_0 A^{(i)}_{N_0}} N_0 A^{(i)}_{N_0}$, we can simulate the value of the rate function $I^h(y) = \sup_{\lambda\in\mathbb{R}}\{\lambda y - \Lambda^h(\lambda)\}$ at $y(\lambda) := (\Lambda^h_{M,N_0})'(\lambda) = \frac{(G_{M,N_0})'(\lambda)}{N_0 G_{M,N_0}(\lambda)}$ by $I^h_{M,N_0}(y(\lambda)) = \lambda y(\lambda) - \Lambda^h_{M,N_0}(\lambda)$. Hence, we have the following algorithm.

Algorithm 7.1

1. Choose the proper step-size h, sample size M, and number of steps N_0, and compute numerical solution $x^{(i)}_n$, $i = 1, 2, \ldots, M, n = 0, 1, \ldots, N_0 - 1$.
2. Set $S_{N_0}(i) = \sum_{n=0}^{N_0-1} x^{(i)}_n$, $i = 1, 2, \ldots, M$.
3. For a given $K > 0$, compute $G_{M,N_0}(\lambda) = \frac{1}{M}\sum_{i=1}^{M} e^{\lambda S^{(i)}_{N_0}}$ and $(G_{M,N_0})'(\lambda) = \frac{1}{M}\sum_{i=1}^{M} S^{(i)}_{N_0} e^{\lambda S^{(i)}_{N_0}}$ for sufficiently many $\lambda \in [-K, K]$.
4. Compute $\Lambda^h_{M,N_0}(\lambda) = \log\left(G_{M,N_0}(\lambda)\right)^{1/N_0}$ and $y(\lambda) = \frac{(G_{M,N_0})'(\lambda)}{N_0 G_{M,N_0}(\lambda)}$.
5. Compute $I^h_{M,N_0}(y(\lambda)) = \lambda y(\lambda) - \Lambda^h_{M,N_0}(\lambda)$ and $I^{h,M,N_0}_{mod}(y(\lambda)) = I^h_{M,N_0}(y(\lambda))/h$.

The numerical realization of the rate functions of $\{B_N\}_{N\geq 1}$ based on a given numerical method is analogous to that of $\{A_N\}_{N\geq 1}$. We refer readers to [20] for more details.

According to Algorithm 7.1, we numerically simulate the modified rate functions of $\{A_N\}_{N\geq 1}$ of the midpoint scheme (symplectic β-method with $\beta = 1/2$) and the



(a) $h = 1.5$

(b) $h = 1$

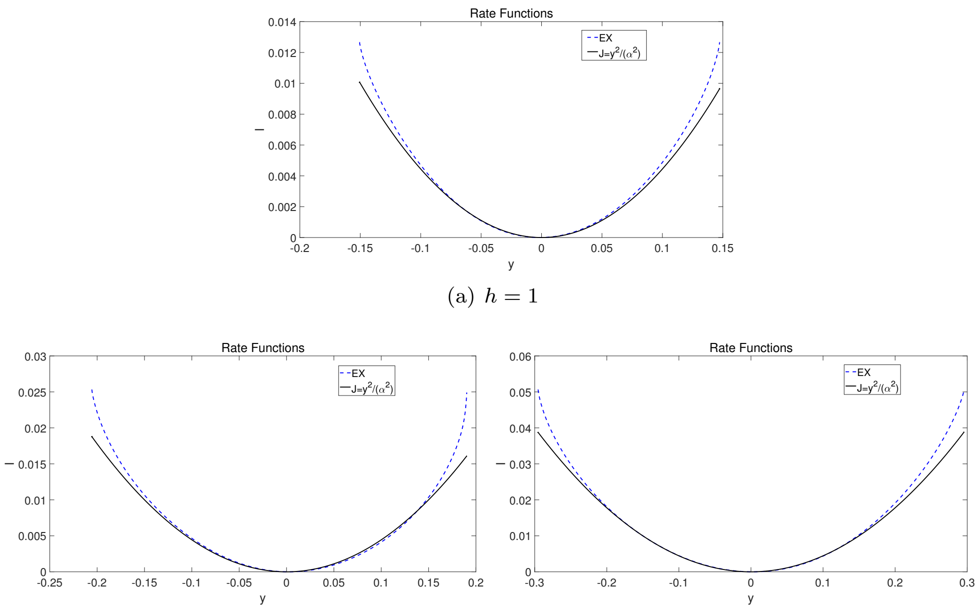
(c) $h = 0.5$

FIG. 1. *Modified rate functions of the midpoint scheme and PEM-MR method under different step-sizes with* $M = 2000$, $N_0 = 600$, $\alpha = 1.5$, $K = 1.5$, *and* $(x_0, y_0) = (0.5, 0)$.

PEM-MR method in section 6. We set initial data $(x_0, y_0) = (0.5, 0)$, $M = 2000$, $N_0 = 600$, and $\alpha = 1.5$. In the third step, G_{M,N_0} is computed at $\lambda(j) = -K + 0.001(j-1)$, $j = 1, 2, \ldots, 2000K + 1$ with $K = 1.5$. It is observed from Figure 1 that as step-size h decreases, the modified rate function $I^h_{mod,1}$ of the midpoint scheme becomes closer to the rate function $I(y) = \frac{y^2}{3\alpha^2}$ of $\{A_T\}_{T>0}$, while the modified rate function of the PEM-MR method gets closer to $I'(y) = \frac{y^2}{2\alpha^2}$. When $h = 1.5, 1, 0.5$, $I^h_{mod,1}$ nearly coincides with $I(y) = \frac{y^2}{3\alpha^2}$ on intervals $[-0.05, 0.05]$, $[-0.15, 0.1]$, and $[-0.35, 0.3]$, respectively. These verify our theoretical results in Theorems 4.3 and 4.4. Analogously, we perform numerical experiments to simulate the modified rate function of the EX method from section 6. As observed in Figure 2, the modified rate function gradually converges to the rate function of $\{B_T\}_{T>0}$, which verifies the result in Theorem 5.2.

Before ending this section, we would like to mention that our theoretical results are meaningful for computing large deviations rate functions. More precisely, if one wants to simulate the rate functions of observables associated with stochastic Hamiltonian systems by computing the logarithmic generating moment function, applying the symplectic method is a prime choice, as shown in Figure 1.

8. Conclusions and future aspects. In this paper, in order to evaluate the ability of the numerical method to preserve the large deviations rate functions associated with the general stochastic Hamiltonian systems, we propose the concept of asymptotical preservation for LDPs. It is shown that stochastic symplectic methods applied to the stochastic test equation, that is, the linear stochastic oscillator, asymptotically preserve the LDPs for $\{A_T\}_{T>0}$ and $\{B_T\}_{T>0}$, but nonsymplectic ones do not. This indicates the superiority of stochastic symplectic methods in the aspect of asymptotically preserving large deviations principles. In fact, there are still many



(a) $h = 1$

(b) $h = 0.5$

(c) $h = 0.25$

FIG. 2. *Modified rate functions of EX method under different step-sizes with* $M = 2000$, $N_0 = 600$, $\alpha = 1.5$, $K = 1.5$, *and* $(x_0, y_0) = (0.5, 0)$.

problems of interest which remain to be solved. We list some possible aspects for future work.

(1) Can the stochastic symplectic methods asymptotically preserve the LDPs for all observables associated with the linear stochastic oscillator?
(2) Can stochastic symplectic methods asymptotically preserve the LDPs for observables associated with the general stochastic Hamiltonian system which is driven by multiplicative noises or in higher dimension?
(3) For a stochastic Hamiltonian partial differential equation which possesses the symplectic or multisymplectic structure, such as the stochastic Schrödinger equation, do the symplectic or multisymplectic numerical methods asymptotically preserve the LDP of the original system?

These problems are very challenging. Because the large deviations rate functions do not generally have explicit expression for more complex SDEs and their numerical solutions, it is difficult to analyze the asymptotical behavior of rate functions of numerical methods. In addition, the large deviations estimates on infinite dimensional Banach spaces are more involved. We leave these problems as the open problems and attempt to study them in our future work.

Appendix A. Proof of Lemma 3.1.

Proof. Using the fact $\sin(n\theta) = \frac{1}{2\boldsymbol{i}}(e^{\boldsymbol{i}n\theta} - e^{-\boldsymbol{i}n\theta})$, one immediately has

$$\sum_{n=1}^{N} \sin(n\theta)a^n = \frac{a\sin(\theta) - a^{N+1}\sin((N+1)\theta) + a^{N+2}\sin(N\theta)}{1 - 2a\cos(\theta) + a^2}.$$

For $a=1$, utilizing the formula $\sin(\alpha)-\sin(\beta)=2\cos(\frac{\alpha+\beta}{2})\sin(\frac{\alpha-\beta}{2})$ gives

$$\sum_{n=1}^{N}\sin(n\theta)=\frac{\sin(\theta)-\sin((N+1)\theta)+\sin(N\theta)}{2(1-\cos(\theta))}=\frac{\cos\left(\frac{\theta}{2}\right)-\cos((N+\frac{1}{2})\theta)}{2\sin\left(\frac{\theta}{2}\right)},$$

which completes the proof. ∎

Appendix B. Proof of Lemma 3.3.

Proof. (1) Assume that $b_1^2+(a_{12}b_2-a_{22}b_1)^2=0$, i.e., $b_1=a_{12}b_2-a_{22}b_1=0$. Noting that $b_1^2+b_2^2\neq 0$, one has $b_2\neq 0$, which leads to $a_{12}=0$. Since $\det(A)=a_{11}a_{22}-a_{12}a_{21}=1$, $a_{11}a_{22}=1$. Hence $a_{11},a_{22}>0$ or $a_{11},a_{22}<0$. It follows from assumptions (A1) and (A2) that $-2<\operatorname{tr}(A)<2$. In this way, $|\operatorname{tr}(A)|=|a_{11}|+|a_{22}|<2$. This is contradictory to $|a_{11}a_{22}|=1$, since $1=\sqrt{|a_{11}a_{22}|}\leq\frac{1}{2}(|a_{11}|+|a_{22}|)<1$. This proves the first conclusion.

(2) Denote $S:=(b_1+a_{12}b_2-a_{22}b_1)^2(4+\operatorname{tr}(A))-2b_1(a_{12}b_2-a_{22}b_1)(2-\operatorname{tr}(A))$, $p:=b_1$, and $q:=a_{12}b_2-a_{22}b_1$. Then $S=(p+q)^2(4+\operatorname{tr}(A))-2pq(2-\operatorname{tr}(A))=\operatorname{tr}(A)((p+q)^2+2pq)+4(p+q)^2-4pq$. By studying the infimum of S in three kinds of cases—$(p+q)^2+2pq>0$, $(p+q)^2+2pq<0$ and $(p+q)^2+2pq=0$—one can prove that $S>0$. ∎

Appendix C. Proof of Theorem 4.1.

Proof. Denote $Z_t=(X_t,Y_t)^\top$, $J=\left(\begin{smallmatrix}0&1\\-1&0\end{smallmatrix}\right)$, $K=\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$. We rewrite (2.1) as $\mathrm{d}Z_t=JZ_t\mathrm{d}t+\alpha K\mathrm{d}W_t$. Let Z be the solution of the above equation at $t+h$, with the deterministic value z at t. Note that for any $u>v\geq 0$, $Z_u=Z_v+J\int_v^u Z_r\mathrm{d}r+\alpha K\int_v^u\mathrm{d}W_r$. Using the above formula, one can show that

$$\begin{aligned}Z&=z+hJz+\alpha K(W_{t+h}-W_t)+J^2\int_t^{t+h}\int_t^s Z_r\mathrm{d}r\mathrm{d}s+\alpha JK\int_t^{t+h}\int_t^s\mathrm{d}W_r\mathrm{d}s\\&=A^{EM}z+\alpha b^{EM}(W_{t+h}-W_t)+R,\end{aligned}\tag{C.1}$$

where $R:=J^2\int_t^{t+h}\int_t^s Z_r\mathrm{d}r\mathrm{d}s+\alpha JK\int_t^{t+h}\int_t^s\mathrm{d}W_r\mathrm{d}s$ with $\|\mathbf{E}R\|_2\leq Ch^2$ and $\mathbf{E}\|R\|_2^2\leq Ch^3$. Further, the one-step approximation based on the method (3.1) is $\widehat{Z}=Az+\alpha b(W_{t+h}-W_t)$. In this way, we obtain

$$\left\|\mathbf{E}(\widehat{Z}-Z)\right\|_2\leq C\left\|A-A^{EM}\right\|_F\|z\|_2+\|\mathbf{E}R\|_2\leq Ch^2,\tag{C.2}$$

where the second equality uses the equivalence of norms in finite dimensional normed linear spaces. In addition, it holds that

$$\mathbf{E}\left\|\widehat{Z}-Z\right\|_2^2\leq C\left\|A-A^{EM}\right\|_F^2\|z\|_2^2+C\alpha^2\left\|b-b^{EM}\right\|_2^2\mathbf{E}(\Delta W^2)+C\mathbf{E}\|R\|_2^2\leq Ch^3.\tag{C.3}$$

It follows from (C.2), (C.3), and [18, Theorem 1.1] that the mean-square convergence order of numerical method (3.1) is at least 1. ∎

Appendix D. Proof of Lemma 4.2.

Proof. If (B) holds, then $a_{11}=1+\mathcal{O}(h^2)$, $a_{22}=1+\mathcal{O}(h^2)$. Thus, $\operatorname{tr}(A)=2+\mathcal{O}(h^2)$, which leads to the assertion (1). Further, $1-\operatorname{tr}(A)+\det(A)=(a_{11}-$

$1)(a_{22}-1)-a_{12}a_{21}$. Noting that $a_{12} \sim h$ and $a_{21} \sim -h$, one has $(1-\operatorname{tr}(A)+\det(A)) \sim h^2$. Finally, since $\lim_{h\to 0} \frac{a_{12}b_2}{h} = \lim_{h\to 0} \frac{(a_{12}-h)(b_2-1)+h(b_2-1)+a_{12}}{h} = 1$, it holds that $\lim_{h\to 0} \frac{b_1+a_{12}b_2-a_{22}b_1}{h} = \lim_{h\to 0} \frac{a_{12}b_2}{h} + \lim_{h\to 0} \frac{b_1(1-a_{22})}{h} = 1$, which is nothing but the assertion (3). ∎

Appendix E. Proof of Lemma 5.1.

Proof. It follows from Lemma 3.3(1) that $b_1^2 + (a_{12}b_2 - a_{22}b_1)^2 \neq 0$. Denote $T = (b_1 + a_{12}b_2 - a_{22}b_1)^2 - b_1(a_{12}b_2 - a_{22}b_1)(2 - \operatorname{tr}(A))$. Then $T = b_1^2 + (a_{12}b_2 - a_{22}b_1)^2 + b_1(a_{12}b_2 - a_{22}b_1)\operatorname{tr}(A)$. Next we show that $T > 0$.

Case 1: $b_1 = 0$ *or* $a_{12}b_2 - a_{22}b_1 = 0$. This associated with $b_1^2 + (a_{12}b_2 - a_{22}b_1)^2 \neq 0$ immediately leads to $T > 0$.

Case 2: $b_1 \neq 0$ *and* $a_{12}b_2 - a_{22}b_1 \neq 0$. We note that under assumptions (A1) and (A2), $-2 < \operatorname{tr}(A) < 2$. If $\operatorname{tr}(A) = 0$, $T > 0$ holds naturally. If $\operatorname{tr}(A) \neq 0$, then $0 < |\operatorname{tr}(A)| < 2$. As a result, $|b_1(a_{12}b_2 - a_{22}b_1)\operatorname{tr}(A)| < 2|b_1(a_{12}b_2 - a_{22}b_1)| \leq b_1^2 + (a_{12}b_2 - a_{22}b_1)^2$, which implies $T > 0$. ∎

Acknowledgments. We would like to thank Prof. Xia Chen (Department of Mathematics, University of Tennessee) for lecturing courses on LDP in the summer of 2018. We are also grateful to the referees for their valuable suggestions.

REFERENCES

[1] A. Abdulle, D. Cohen, G. Vilmart, and K. Zygalakis, *High weak order methods for stochastic differential equations based on modified equations*, SIAM J. Sci. Comput., 34 (2012), pp. A1800–A1823, https://doi.org/10.1137/110846609.

[2] C. Anton, *Weak backward error analysis for stochastic Hamiltonian Systems*, BIT, 59 (2019), pp. 613–646, https://doi.org/10.1007/s10543-019-00747-6.

[3] C. Chen and J. Hong, *Symplectic Runge-Kutta semidiscretization for stochastic Schrödinger equation*, SIAM J. Numer. Anal., 54 (2016), pp. 2569–2593, https://doi.org/10.1137/151005208.

[4] C. Chen, J. Hong, and C. Huang, *Stochastic Modified Equations for Symplectic Methods Applied to Rough Hamiltonian Systems Based on the Wong–Zakai Approximation*, arXiv:1907.02825.

[5] C. Chen, J. Hong, and L. Ji, *Mean-square convergence of a symplectic local discontinuous Galerkin method applied to stochastic linear Schrödinger equation*, IMA J. Numer. Anal., 37 (2017), pp. 1041–1065, https://doi.org/10.1093/imanum/drw023.

[6] C. Chen, J. Hong, and L. Zhang, *Preservation of physical properties of stochastic Maxwell equations with additive noise via stochastic multi-symplectic methods*, J. Comput. Phys., 306 (2016), pp. 500–519, https://doi.org/10.1016/j.jcp.2015.11.052.

[7] X. Chen, *Random walk intersections. Large Deviations and Related Topics*, Math. Surveys Monogr. 157, American Mathematical Society, Providence, RI, 2010, https://doi.org/10.1090/surv/157.

[8] D. Cohen, *On the numerical discretisation of stochastic oscillators*, Math. Comput. Simulation, 82 (2012), pp. 1478–1495, https://doi.org/10.1016/j.matcom.2012.02.004.

[9] J. Cui, J. Hong, Z. Liu, and W. Zhou, *Stochastic symplectic and multi-symplectic methods for nonlinear Schrödinger equation with white noise dispersion*, J. Comput. Phys., 342 (2017), pp. 267–285, https://doi.org/10.1016/j.jcp.2017.04.029.

[10] A. Debussche and E. Faou, *Weak backward error analysis for SDEs*, SIAM J. Numer. Anal., 50 (2012), pp. 1735–1752, https://doi.org/10.1137/110831544.

[11] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Stoch. Model. Appl. Probab., Springer-Verlag, Berlin, 2010, https://doi.org/10.1007/978-3-642-03311-7.

[12] J. Deng, C. Anton, and Y. Wong, *High-order symplectic schemes for stochastic Hamiltonian systems*, Commun. Comput. Phys., 16 (2014), pp. 169–200, https://doi.org/10.4208/cicp.311012.191113a.

[13] G. Ferré and H. Touchette, *Adaptive sampling of large deviations*, J. Stat. Phys., 172 (2018), pp. 1525–1544, https://doi.org/10.1007/s10955-018-2108-8.

[14] J. Hong, L. Sun, and X. Wang, *High order conformal symplectic and ergodic schemes for the stochastic Langevin equation via generating functions*, SIAM J. Numer. Anal., 55 (2017), pp. 3006–3029, https://doi.org/10.1137/17M111691X.

[15] J. Hong and X. Wang, *Invariant Measures for Stochastic Nonlinear Schrödinger Equations. Numerical Approximations and Symplectic Structures*, Lecture Notes in Math. 2251, Springer-Verlag, Berlin, 2019.

[16] A. Klenke, *Probability Theory. A Comprehensive Course*, Universitext, Springer-Verlag, Berlin, 2008, https://doi.org/10.1007/978-1-84800-048-3.

[17] X. Mao, *Stochastic Differential Equations and Applications*, 2nd ed., Horwood Publishing, Chichester, UK, 2008, https://doi.org/10.1533/9780857099402.

[18] G. Milstein and M. Tretyakov, *Stochastic Numerics for Mathematical Physics*, Sci. Comput., Springer-Verlag, Berlin, 2004, https://doi.org/10.1007/978-3-662-10063-9.

[19] E. Palamarchuk, *An analytic study of the Ornstein–Uhlenbeck process with time-varying coefficients in the modeling of anomalous diffusions*, Autom. Remote Comtrol, 79 (2018), pp. 289–299, https://doi.org/10.1134/S000511791802008X.

[20] C. M. Rohwer, F. Angeletti, and H. Touchette, *Convergence of large-deviation estimators*, Phys. Rev. E, 92 (2015), 052104, https://doi.org/10.1103/PhysRevE.92.052104.

[21] M. Senosiain and A. Tocino, *A review on numerical schemes for solving a linear stochastic oscillator*, BIT, 55 (2015), pp. 515–529, https://doi.org/10.1007/s10543-014-0507-z.

[22] T. Shardlow, *Modified equations for stochastic differential equations*, BIT, 46 (2006), pp. 111–125, https://doi.org/10.1007/s10543-005-0041-0.

[23] L. Wang and J. Hong, *Generating functions for stochastic symplectic methods*, Discrete Contin. Dyn. Syst., 34 (2014), pp. 1211–1228, https://doi.org/10.3934/dcds.2014.34.1211.

[24] L. Wang, J. Hong, and L. Sun, *Modified equations for weakly convergent stochastic symplectic schemes via their generating functions*, BIT, 56 (2016), pp. 1131–1162, https://doi.org/10.1007/s10543-015-0583-8.

[25] K. Zygalakis, *On the existence and the applications of modified equations for stochastic differential equations*, SIAM J. Sci. Comput., 33 (2011), pp. 102–130, https://doi.org/10.1137/090762336.