Exact Calculation of Quantum Thermal Average from Continuous Loop Path Integral Molecular Dynamics

Xuda Ye* Zhennan Zhou†

July 11, 2023

Abstract

The quantum thermal average plays a central role in describing the thermodynamic properties of a quantum system. From the computational perspective, the quantum thermal average can be computed by the path integral molecular dynamics (PIMD), but the knowledge on the quantitative convergence of such approximations is lacking. We propose an alternative computational framework named the continuous loop path integral molecular dynamics (CL-PIMD), which replaces the ring polymer beads by a continuous loop in the spirit of the Feynman–Kac formula. By truncating the number of normal modes to a finite integer $N \in \mathbb{N}$, we quantify the discrepency of the statistical average of the truncated CL-PIMD from the true quantum thermal average, and prove that the truncated CL-PIMD has uniform-in-N geometric ergodicity. These results show that the CL-PIMD provides an accurate approximation to the quantum thermal average, and serves as a mathematical justification of the PIMD methodology.

Keywords quantum thermal average, path integral molecular dynamics, Feynman–Kac formula, geometric ergodicity, generalized Γ calculus

AMS subject classifications 37A30, 82B31, 81S40

1 Introduction

Calculation of the quantum thermal average plays an important role in quantum physics and quantum chemistry, not only because it fully characterizes the canonical ensemble of the quantum system, but also because of its wide applications in describing the thermal properties of complex quantum systems, including the idea quantum gases [1], chemical reaction rates [2, 3], density of states of crystals [4], quantum phase transitions [5], etc. However, since the computational cost of direct discretization methods (finite difference, pseudospectral methods, etc.) grows exponentially with the spatial dimension [6, 7], the exact calculation of the quantum thermal average is hardly affordable. In the past decades, there have been numerous methods committed to compute the quantum thermal average approximately, and the path integral molecular dynamics (PIMD) is among the most prevailing ones.

*Beijing International Center for Mathematical Research, Peking University, Beijing, 100871, P. R. China. Email: abneryepku@pku.edu.cn

†Beijing International Center for Mathematical Research, Peking University, Beijing, 100871, P. R. China. Email: zhennan@bicmr.pku.edu.cn

The PIMD is a computational framework to obtain accurate quantum thermal averages. Using the imaginary time-slicing in Feynman's path integral theory, the PIMD maps the quantum system in $\mathbb{R}^d$ to a ring polymer of P beads in $\mathbb{R}^{dP}$, where each bead represents a classical duplicate of the original quantum system, and adjacent beads are connected by a harmonic oscillator. When the number of beads P is large enough, the classical Boltzmann distribution of the ring polymer in $\mathbb{R}^{dP}$ is expected to yield an accurate approximation of the quantum thermal average. Since the development of the PIMD in 1970s, this framework has been widely used in the calculations in the chemical reaction rates [8–10], transition state theory [11] and tunneling splittings [12, 13]. The variants of the PIMD, the ring polymer molecular dynamics (RPMD) [14,15] and the centroid molecular dynamics (CMD) [16, 17], are employed to compute the quantum correlation function. Recently, there have also been fruitful studies on designing efficient and accurate algorithms to enhance the numerical performance of the standard PIMD [18–21]. The procedure to compute the quantum thermal average in the PIMD framework is demonstrated as follows:

1. **Ring polymer approximation**. Pick an integer $P \in \mathbb{N}$ and map the quantum system to a ring polymer system of P beads, where each bead is a classical duplicate.
2. **Construction of sampling**. Construct an ergodic stochastic process (e.g., Langevin) whose invariant distribution is the Boltzmann distribution of the ring polymer system.
3. **Stochastic simulation**. Evolve the stochastic process with a time discretization method, and approximate the quantum thermal average by the time average of stochastic process in the long-time simulation.

Although the PIMD has become a mature framework to compute the quantum thermal average, its mathematical understanding is still widely lacking. There are two natural questions we can ask about the ergodic stochastic process which samples the Boltzmann distribution of the ring polymer system. (a) Why does the distribution of the ring polymer reproduce the quantum thermal average as the number of beads $P \to \infty$? (b) Does the ergodic stochastic process has a convergence rate which does not depend on P? These questions are in fact related to the convergence properties of the ergodic stochastic processes in the PIMD. Although these properties have been partially validated when the potential function is quadratic [22, 23], a rigorous justification of the convergence of the PIMD is still to be investigated.

The convergence analysis of the PIMD has its unique significance among the large variety of numerical methods to compute the quantum thermal average. Also, the computational cost of the PIMD does not grow exponentially with the spatial dimensional d, thus it suits high dimensional simulations. To evaluate the efficiency of the PIMD, it is vital to estimate the bias and convergence rate of the ergodic stochastic process, which are essentially the questions (a)(b) proposed before. Therefore, the convergence analysis of the PIMD makes up the mathematical foundation for the numerical calculation of the quantum thermal average in high dimensions.

Despite of the importance of the convergence analysis of the PIMD, estimating either its bias or its convergence rate towards the invariant distribution is difficult. Instead of studying the standard PIMD, we propose a novel computational framework, the continuous loop path integral molecular dynamics (CL-PIMD), for the accurate calculation of the quantum thermal average. In this paper, we derive the CL-PIMD from the infinite bead limit of the standard PIMD using the Feynman–Kac formula [24], and present a quantitative convergence analysis of the CL-PIMD. In particular, the counterparts of questions (a)(b) for the CL-PIMD both have positive answers. In this way, the CL-PIMD can be viewed as an alternative to the standard PIMD, while the sampling efficiency of the CL-PIMD can be explicitly estimated.

It is worth emphasizing that the continuous loop in the CL-PIMD is represented in the normal mode coordinates [25] rather than the spatial coordinates of the ring polymer beads. Using the normal modes allows us to derive the infinite bead limit of the PIMD in a more convenient way. From the computational viewpoint, we need to truncate the number of normal modes in the continuous loop to a finite integer $N \in \mathbb{N}$ to implement the CL-PIMD numerically. The resulting dynamics is named as the truncated CL-PIMD, which is a finite-dimensional underdamped Langevin dynamics in $\mathbb{R}^{dN}$, and can be implemented with a suitable time discretization scheme. The convergence analysis of the CL-PIMD mentioned earlier is actually referring to (a) the quantification of the difference between the statistical average of the truncated CL-PIMD and the true quantum thermal average, and (b) the uniform-in-N geometric ergodicity of the truncated CL-PIMD, i.e., the convergence rate of the CL-PIMD towards the invariant distribution does not depend on the normal modes N.

We briefly introduce the strategies to prove the convergence of the CL-PIMD. For the question (a), we show that the statistical average of the truncated CL-PIMD is a Cauchy sequence as the number of normal modes $N \to \infty$, then employ the Trotter product formula [24] to show that the limit of the statistical average coincides with the quantum thermal average. For the question (b), we utilize the log-Sobolev inequality in the Bakry–Émery theory [26] to establish the uniform-in-N ergodicity of the overdamped Langevin dynamics, then apply the generalized Γ calculus [27, 28] to prove the uniform-in-N ergodicity of the underdamped Langevin dynamics.

The paper is organized as follows. In Section 2 we employ the normal mode coordinates to derive the CL-PIMD from the infinite bead limit of the standard PIMD in a formal regime. In Section 3 we establish the rigorous mathematical convergence theory for the truncated CL-PIMD. In Section 4 we demonstrate the numerical method to discretize the CL-PIMD. In Section 5 we implement the numerical experiments to show the performance of the CL-PIMD in computing the quantum thermal average.

2 Formal derivation of CL-PIMD from infinite bead limit

In this section we derive the formulation of the CL-PIMD from the infinite bead limit of the standard PIMD. We first review the standard path integral representation with a finite number of beads, then obtain the energy function of the continuous loop path integral representation. Finally, we represent the energy function in the normal mode coordinates and construct the infinite-dimensional underdamped Langevin dynamics corresponding to the energy function.

2.1 Review of the standard path integral representation

In this paper, we consider the quantum system in $\mathbb{R}^d$ given by the Hamiltonian operator

$$\hat{H} = \frac{\hat{p}^2}{2} + V(\hat{q}), \tag{2.1}$$

where $\hat{q}$ and $\hat{p}$ are the position and momentum operators in $\mathbb{R}^d$, and $V(\cdot)$ is a real-valued potential function in $\mathbb{R}^d$. When the quantum system is at a constant temperature $T = 1/\beta$, the state of the system can be described by the canonical ensemble with the density operator $e^{-\beta\hat{H}}$, and thus the partition function is given by $\mathcal{Z} = \mathrm{Tr}[e^{-\beta\hat{H}}]$. The quantum thermal average of the system means the average of an observable operator $O(\hat{q})$ in the canonical ensemble $e^{-\beta\hat{H}}$, i.e.,

$$\langle O(\hat{q})\rangle_\beta = \frac{\mathcal{Z}_O}{\mathcal{Z}} = \frac{\mathrm{Tr}[e^{-\beta\hat{H}}O(\hat{q})]}{\mathrm{Tr}[e^{-\beta\hat{H}}]}, \tag{2.2}$$

where $\mathcal{Z}_O = \mathrm{Tr}[e^{-\beta\hat{H}}O(\hat{q})]$ is the partition function involving the observable operator $O(\hat{q})$. Here, we assume the observable operator $O(\hat{q})$ depends only on the position operator $\hat{q}$, where $O(\cdot)$ is a real-valued function in $\mathbb{R}^d$.

The starting point of the PIMD is to replace the quantum system by a ring polymer of P beads, where each bead reprents a classical replica of the original system [29,30]. By inserting the orthonormal expansion $I_d = \int_{\mathbb{R}^d} |x_j\rangle\langle x_j|$ in the quantum thermal average (2.2), we can apply the Trotter splitting

$$e^{-\beta_P\hat{H}} = e^{-\frac{\beta_P}{2}V(\hat{q})}e^{-\frac{\beta_P}{2}\hat{p}^2}e^{-\frac{\beta_P}{2}V(\hat{q})} + O(\beta_P^3), \quad \beta_P = \beta/P, \tag{2.3}$$

to approximate the propagator $\langle x_j|e^{-\beta_P\hat{H}}|x_{j+1}\rangle$ for $j = 1, \cdots, P$. Finally, the partition functions Z and Z_O are approximated as

$$\mathcal{Z} = \lim_{P\to\infty} \frac{1}{(2\pi\beta_P)^{\frac{dP}{2}}} \int_{\mathbb{R}^{dP}} \exp\bigg(-\frac{1}{2\beta_P}\sum_{j=1}^P |x_j - x_{j+1}|^2\bigg) \times \Theta(x_1, \cdots, x_P) \times \mathrm{d}x_1\cdots\mathrm{d}x_P, \tag{2.4a}$$

$$\mathcal{Z}_O = \lim_{P\to\infty} \frac{1}{(2\pi\beta_P)^{\frac{dP}{2}}} \int_{\mathbb{R}^{dP}} \exp\bigg(-\frac{1}{2\beta_P}\sum_{j=1}^P |x_j - x_{j+1}|^2\bigg) \times \Theta_O(x_1, \cdots, x_P) \times \mathrm{d}x_1\cdots\mathrm{d}x_P, \tag{2.4b}$$

where the functions $\Theta(x_1, \cdots, x_P)$ and $\Theta_O(x_1, \cdots, x_P)$ are given by

$$\Theta(x_1, \cdots, x_P) = \prod_{j=1}^P e^{-\beta_P V(x_j)}, \quad \Theta_O(x_1, \cdots, x_P) = \prod_{j=1}^P e^{-\beta_P V(x_j)} \times \bigg(\frac{1}{P}\sum_{j=1}^P O(x_j)\bigg), \tag{2.5}$$

and the convention $x_1 = x_{P+1}$ is assumed. The summation $\sum_{j=1}^P |x_j - x_{j+1}|^2$ represents the internal potential function of the ring polymer, where the adjacent beads are connected by the harmonic spring potential. By defining the energy function

$$\mathcal{E}_P(x) = \frac{1}{2\beta_P}\sum_{j=1}^P |x_j - x_{j+1}|^2 + \beta_P\sum_{j=1}^P V(x_j), \quad x \in \mathbb{R}^{dP}, \tag{2.6}$$

we identify the ring polymer distribution as $\pi_P(x) \propto \exp(-\mathcal{E}_P(x))$, then the thermal average $\langle O(\hat{q})\rangle_\beta$ is expected to the limit of the statistical average

$$\langle O(\hat{q})\rangle_\beta = \lim_{P\to\infty} \int_{\mathbb{R}^{dP}} \bigg(\frac{1}{P}\sum_{j=1}^P O(x_j)\bigg)\pi_P(x)\mathrm{d}x. \tag{2.7}$$

Therefore, to compute the quantum thermal average $\langle O(\hat{q})\rangle_\beta$ numerically, we only need to pick a large integer $P \in \mathbb{N}$ and sample the probability distribution $\pi_P(x)$ with a suitable stochastic sampling method, e.g., the Langevin dynamics.

2.2 Path integral representation in normal mode coordinates

The idea behind the CL-PIMD is to observe that the P beads $x_1, \cdots, x_P \in \mathbb{R}^d$ in the ring polymer distribution $\pi_P(x)$ can be viewed as a continuous loop $x(\cdot) \in C([0,\beta];\mathbb{R}^d)$, where the P grid points $\{x(j\beta_P)\}_{j=1}^P$ correspond to P beads $\{x_j\}_{j=1}^P$. Figure 1 shows an instance of the continuous loop.

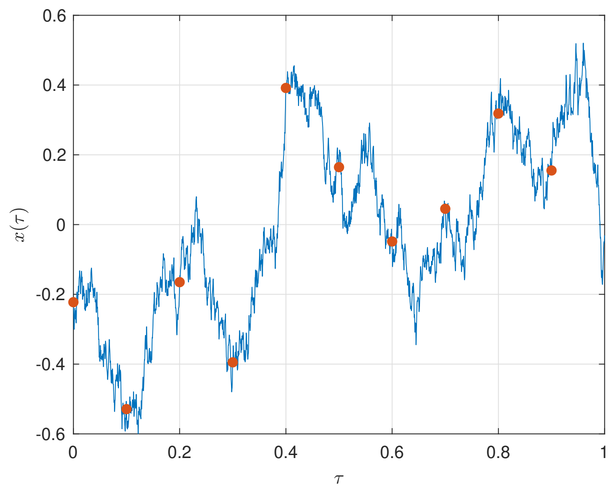

Figure 1: Continuous loop $x(\tau)$ with dimension $d = 1$. The 10 bead positions are marked in red.

mode	ξ_0	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5	ξ_6
value	-0.1242	0.0841	-0.0595	-0.1774	-0.0954	0.0565	-0.1184

Table 1: Normal mode coordinates $\{\xi_k\}_{k=0}^6$ of the continuous loop $x(\tau)$ shown in Figure 1.

For a given continuous loop $x(\cdot) \in C([0,\beta];\mathbb{R}^d)$, if we set $x_j = x(j\beta_P)$ in the energy function $\mathcal{E}_P(x)$ (2.6), then formally we have

$$x\big((j+1)\beta_P\big) - x(j\beta_P) \approx \beta_P x'(j\beta_P), \qquad j = 1, \cdots, P, \tag{2.8}$$

where the convention $x(\tau) = x(\tau+\beta)$ is assumed. Then in the limit $P \to \infty$ we have formally

$$\lim_{P\to\infty} \mathcal{E}_P(x) \approx \frac{1}{2}\int_0^\beta |x'(\tau)|^2 \mathrm{d}\tau + \int_0^\beta V(x(\tau))\mathrm{d}\tau. \tag{2.9}$$

For a given continuous loop $x(\cdot) \in C([0,\beta];\mathbb{R}^d)$, it is natural to define the energy function $\mathcal{E}(x)$ by

$$\mathcal{E}(x) = \frac{1}{2}\int_0^\beta |x'(\tau)|^2 \mathrm{d}\tau + \int_0^\beta V(x(\tau))\mathrm{d}\tau, \tag{2.10}$$

where the integral of $|x'(\cdot)|^2$ contributes to the ring polymer potential, while the integral of $V(x(\cdot))$ contributes to the potential V acting on the whole loop $\{x(\tau)\}_{\tau\in[0,\beta]}$. In this way, we expect to sample the Boltzmann distribution formally written as $\pi(x) \propto \exp(-\mathcal{E}(x))$ to obtain the accurate quantum thermal average $\langle O(\hat{q})\rangle_\beta$.

Although the energy function (2.10) has been well-known in the study of the Feynman–Kac formula [31], it is rarely discussed how to use (2.10) itself rather than its ring polymer approximation to design a proper stochastic method to compute the thermal average. In recent years, the authors of [22, 32] have studied the infinite bead limit of the standard PIMD, and identify the limit of the ring polymer beads $x_1, \cdots, x_P \in \mathbb{R}^d$ as a continuous loop $x(\cdot)$ in $C([0,\beta];\mathbb{R}^d)$. In particular, a stochastic partial differential equation (SPDE) is formally proposed in [22] to characterize the limit dynamics of the ring polymer beads, and it is proved in [32] that the Hamiltonian Monte Carlo method for the standard path integral representation has dimension-free ergodicity. When the potential function $V(q)$ is harmonic, the authors of [23,33] construct a strongly stable numerical method for the underdamped Langevin dynamics with dimension-free ergodicity. All these results support that the PIMD has a mathematically well-defined limit as the number of beads $P \to \infty$, but how to define such a limit stochastic process in a rigorous regime is still a difficult question.

Using the similar approach in [32], we identify the limit of the ring polymer beads $x_1, \cdots, x_P \in \mathbb{R}^d$ as a continuous loop $x(\cdot)$ which lives in the Hilbert space

$$\mathbb{H} = L_p^2([0,\beta];\mathbb{R}^d), \qquad (x,\tilde{x})_{\mathbb{H}} := \int_0^\beta x(\tau)\tilde{x}(\tau)\mathrm{d}\tau, \qquad x,\tilde{x} \in \mathbb{H}, \tag{2.11}$$

where the subscript p means the periodic boundary condition, i.e., $x(0) = x(\beta)$.

Remark We use the Hilbert space $\mathbb{H} = L_p^2([0,\beta];\mathbb{R}^d)$ rather than the continuous function space $C([0,\beta];\mathbb{R}^d)$ because it is convenient to expand $x(\cdot) \in \mathbb{H}$ in the normal mode coordinates.

As the number of beads P tends to infinity, the limit of the partition functions $\mathcal{Z}$ and $\mathcal{Z}_O$ in (2.4) are formally given by

$$\mathcal{Z} = \int_{\mathbb{H}} \exp\left(-\frac{1}{2}\int_0^\beta |x'(\tau)|^2\mathrm{d}\tau\right)\Theta(x)\,\mathrm{d}x, \quad \mathcal{Z}_O = \int_{\mathbb{H}} \exp\left(-\frac{1}{2}\int_0^\beta |x'(\tau)|^2\mathrm{d}\tau\right)\Theta_O(x)\,\mathrm{d}x, \tag{2.12}$$

where $\Theta(x)$ and $\Theta_O(x)$ are functionals on $\mathcal{H}$ defined by

$$\Theta(x) = \exp\left(-\int_0^\beta V(x(\tau))\mathrm{d}\tau\right), \quad \Theta_O(x) = \exp\left(-\int_0^\beta V(x(\tau))\mathrm{d}\tau\right)\times\left(\frac{1}{\beta}\int_0^\beta O(x(\tau))\mathrm{d}\tau\right). \tag{2.13}$$

It should be emphasized the integrals in (2.12) are merely formal and do have a mathematically rigorous definition, because there is not flat distribution in $\mathbb{H}$.

The continuous loop $x(\cdot)$ is a continuous function in $[0,\beta]$, and sampling such loops directly in $\mathbb{H}$ can be difficult. Benefited from the separability of the Hilbert space $\mathbb{H}$, we can represent $x(\cdot)$ and the energy function $\mathcal{E}(x)$ in the normal mode coordinates. Consider the following eigenvalue problem of the Laplace operator with the periodic boundary condition

$$-\ddot{c}_k(\tau) = \omega_k^2 c_k(\tau), \qquad k = 0,1,2,\cdots, \tag{2.14}$$

where $\omega_k \geqslant 0$ and $c_k(\cdot) \in \mathbb{H}$. The eigenvalues and eigenfunctions are explicitly given by

$$\begin{aligned}
\omega_0 &= 0, & c_0(\tau) &= \sqrt{\frac{1}{\beta}}; \\
\omega_{2k-1} &= \frac{2k\pi}{\beta} & c_{2k-1}(\tau) &= \sqrt{\frac{2}{\beta}} \sin\left(\frac{2k\pi\tau}{\beta}\right), \quad k \in \mathbb{N}; \\
\omega_{2k} &= \frac{2k\pi}{\beta}, & c_{2k}(\tau) &= \sqrt{\frac{2}{\beta}} \cos\left(\frac{2k\pi\tau}{\beta}\right), \quad k \in \mathbb{N},
\end{aligned}$$

which is exactly the orthonormal Fourier basis in $\mathbb{H}$. As a consequence, any continuous loop $x(\cdot) \in \mathbb{H}$ can be uniquely represented as

$$x(\tau) = \sum_{k=0}^{\infty} \xi_k c_k(\tau), \qquad \tau \in [0, \beta], \tag{2.15}$$

where $\{\xi_k\}_{k=0}^{\infty}$ in $\mathbb{R}^d$ are the coordinates of $x(\cdot) \in \mathbb{H}$ in different normal modes. For this reason, $\{\xi_k\}_{k=0}^{\infty}$ are referred to as the normal mode coordinates in this paper. Conversely, for a given continuous loop $x(\cdot) \in \mathbb{H}$, the normal mode coordinates are given by

$$\xi_k = \int_0^\beta x(\tau) c_k(\tau) \mathrm{d}\tau, \qquad k = 0, 1, 2, \cdots. \tag{2.16}$$

Table 1 shows the first few normal mode coordinates of the continuous loop shown in Figure 1.

In a formal manner, the energy function $\mathcal{E}(x)$ given in (2.10) can be written in terms of the normal mode coordinates $\xi = \{\xi_k\}_{k=0}^{\infty}$ in $\mathbb{R}^d$:

$$\begin{aligned}
\mathcal{E}(\xi) &= \frac{1}{2}(\dot{x}, \dot{x})_{\mathbb{H}} + \int_0^\beta V(x(\tau))\mathrm{d}\tau = \frac{1}{2}(x, -\ddot{x})_{\mathbb{H}} + \int_0^\beta V(x(\tau))\mathrm{d}\tau \\
&= \frac{1}{2}\left(\sum_{k=0}^{\infty} \xi_k c_k(\tau), \sum_{k=0}^{\infty} \omega_k^2 \xi_k c_k(\tau)\right)_{\mathbb{H}} + \int_0^\beta V\left(\sum_{k=0}^{\infty} \xi_k c_k(\tau)\right)\mathrm{d}\tau \\
&= \frac{1}{2}\sum_{k=0}^{\infty} \omega_k^2 |\xi_k|^2 + \int_0^\beta V\left(\sum_{k=0}^{\infty} \xi_k c_k(\tau)\right)\mathrm{d}\tau.
\end{aligned} \tag{2.17}$$

Therefore, the partition functions $\mathcal{Z}$ and $\mathcal{Z}_O$ can be formally written as

$$\mathcal{Z} = \int \exp\left(-\frac{1}{2}\sum_{k=0}^{\infty} \omega_k^2 |\xi_k|^2\right) \Theta(\xi)\,\mathrm{d}\xi, \quad \mathcal{Z}_O = \int \exp\left(-\frac{1}{2}\sum_{k=0}^{\infty} \omega_k^2 |\xi_k|^2\right) \Theta_O(\xi)\,\mathrm{d}\xi, \tag{2.18}$$

where the functions $\Theta(\xi)$ and $\Theta_O(\xi)$ are represented in the normal mode coordinates

$$\Theta(\xi) = \exp\left(-\int_0^\beta V\left(\sum_{k=0}^{\infty} \xi_k c_k(\tau)\right)\mathrm{d}\tau\right), \tag{2.19a}$$

$$\Theta_O(\xi) = \exp\left(-\int_0^\beta V\left(\sum_{k=0}^{\infty} \xi_k c_k(\tau)\right)\mathrm{d}\tau\right) \times \left(\frac{1}{\beta}\int_0^\beta O\left(\sum_{k=0}^{\infty} \xi_k c_k(\tau)\right)\mathrm{d}\tau\right). \tag{2.19b}$$

Compared to standard path integral representation in (2.4), the integration in (2.18) involves countable number of variables $\{\xi_k\}_{k=0}^{\infty}$ rather than the whole continuous loop $x(\cdot) \in \mathbb{H}$.

2.3 Derivation of continuous loop path integral molecular dynamics

In order to caculate the quantum thermal average $\langle O(\hat{q})\rangle_\beta = \mathcal{Z}_O/\mathcal{Z}$, introduce the energy function

$$\mathcal{E}(\xi) = \frac{1}{2}\sum_{k=0}^\infty \omega_k^2|\xi_k|^2 + \int_0^\beta V\bigg(\sum_{k=0}^\infty \xi_k c_k(\tau)\bigg)\mathrm{d}\tau. \tag{2.20}$$

Formally, the Boltzmann distribution corresponding to the energy function $\mathcal{E}(\xi)$ can be written as $\pi(\xi) \propto \exp(-\mathcal{E}(\xi))$. Although the definition of $\pi(\xi)$ here is still formal, we can still construct an infinite-dimensional stochastic process which formally preserves $\pi(\xi)$ as the invariant distribution. Introduce the constant $a > 0$ and rewrite the energy function $E(\xi)$ as

$$\mathcal{E}(\xi) = \frac{1}{2}\sum_{k=0}^\infty (\omega_k^2 + a^2)|\xi_k|^2 + \int_0^\beta V^a\bigg(\sum_{k=0}^\infty \xi_k c_k(\tau)\bigg)\mathrm{d}\tau, \tag{2.21}$$

where the potential function $V^a(q) := V(q) - a^2|q|^2/2$. For the sake of efficient sampling, we introduce the auxillary velocity variables $\{\eta_k\}_{k=0}^\infty$ to obtain the Hamiltonian function

$$\mathcal{H}(\xi,\eta) = \frac{1}{2}\sum_{k=0}^\infty (\omega_k^2 + a^2)(|\xi_k|^2 + |\eta_k|^2) + \int_0^\beta V^a\bigg(\sum_{k=0}^\infty \xi_k c_k(\tau)\bigg)\mathrm{d}\tau. \tag{2.22}$$

Here, we choose the coefficients of each ξ_k and η_k to be the same so that the Hamiltonian trajectory of (ξ_k, η_k) is close to a circle, which is convenient to design the numerical method. The rapid growth of ω_k as k tends to infinity is actually known as the stiffness in the PIMD, and our choice of the coefficients $\omega_k^2+a^2$ is inspired from the preconditioning method [22,32], which is a common technique in resolving the stiffness of the PIMD.

The gradient of $\mathcal{H}(\xi,\eta)$ with respect to each ξ_k and η_k is

$$\frac{\partial\mathcal{H}}{\partial\xi_k} = (\omega_k^2 + a^2)\xi_k + \int_0^\beta \nabla V^a(x(\tau))c_k(\tau)\mathrm{d}\tau, \qquad \frac{\partial\mathcal{H}}{\partial\eta_k} = (\omega_k^2 + a^2)\eta_k, \tag{2.23}$$

where $x(\tau) \in \mathbb{H}$ is the continuous loop given in (2.15). Then the underdamped Langevin dynamics corresponding to the Hamiltonian $\mathcal{H}(\xi,\eta)$ is given by

$$\begin{cases} \dot{\xi}_k = \dfrac{1}{\omega_k^2 + a^2}\dfrac{\partial\mathcal{H}}{\partial\eta_k}, \\ \dot{\eta}_k = -\dfrac{1}{\omega_k^2 + a^2}\dfrac{\partial\mathcal{H}}{\partial\xi_k} - \gamma_k\eta_k + \sqrt{\dfrac{2\gamma_k}{\omega_k^2 + a^2}}\dot{B}_k, \end{cases} \qquad k = 0,1,2,\cdots, \tag{2.24}$$

where $\{\gamma_k\}_{k=0}^\infty$ are the damping rates and $\{B_k\}_{k=0}^\infty$ are independent Brownian motions in $\mathbb{R}^d$. Inserting the expressions of $\partial H/\partial\xi_k$ and $\partial\mathcal{H}/\partial\eta_k$ to the Langevin dynamics (2.24), we obtain

$$\begin{cases} \dot{\xi}_k = \eta_k, \\ \dot{\eta}_k = -\xi_k - \dfrac{1}{\omega_k^2 + a^2}\displaystyle\int_0^\beta \nabla V^a(x(\tau))c_k(\tau)\mathrm{d}\tau - \gamma_k\eta_k + \sqrt{\dfrac{2\gamma_k}{\omega_k^2 + a^2}}\dot{B}_k, \end{cases} \qquad k = 0,1,2,\cdots. \tag{2.25}$$

The underdamped Langevin dynamics (2.25) is referred to as the continuous loop path integral molecular dynamics (CL-PIMD) thereafter. Note that the invariant distribution of (2.25) is formally $\exp(-\mathcal{H}(\xi,\eta))$, whose marginal distribution in ξ is exactly $\pi(\xi) \propto \exp(-\mathcal{E}(\xi))$.

Although (2.25) mainly involves the normal mode coordinates $\{\xi_k\}_{k=0}^\infty$ and $\{\eta_k\}_{k=0}^\infty$, we note that (2.25) is essentially describing the motion of the continuous loop $x(\cdot) \in \mathbb{H}$. In the evolution of (2.25), the normal mode coordinates $\{\xi_k\}_{k=0}^\infty$ give rise to the continuous loop $x(\cdot)$ through the normal mode expansion (2.15). As long as the potential function $V^a(q)$ is nonzero, the drift force even on a single normal mode

$$\frac{1}{\omega_k^2 + a^2} \int_0^\beta \nabla V^a(x(\tau)) c_k(\tau), \qquad k = 0, 1, 2, \cdots \tag{2.26}$$

can involve all the normal mode coordinates $\{\xi_k\}_{k=0}^\infty$. Therefore, the evolution of $\{\xi_k\}_{k=0}^\infty$ is governed by the global information of the continuous loop $\{x(\tau) : \tau \in [0,\beta]\}$, and the formulation of (2.25) is merely the projection of the dynamics of $x(\cdot) \in \mathbb{H}$ onto each normal mode.

Note that the derivation of the CL-PIMD (2.25) is completely formal, and either its existence or well-posedness is not justified. Since (2.25) is an infinite-dimensional Langevin dynamics, it is not only difficult to validate the rationality of (2.25), but also rarely possible to construct the time discretization scheme which can be implemented numerically. Therefore, we choose to truncate the number of normal modes in the CL-PIMD (2.25) to a finite integer $N \in \mathbb{N}$, and study the ergodic properties of truncated CL-PIMD instead. In Section 3, we validate the convergence properties of the truncated CL-PIMD, and show that the quantum thermal average can be approximated by the invariant distribution of the truncated CL-PIMD in the limit $N \to \infty$.

Remark Here are some remarks on the CL-PIMD (2.25).

1. The idea of the CL-PIMD (2.25) is very similar to the Matsubara dynamics [34–36]. Both dynamics employ lowest few normal modes as the sampling variables in the continuous loop limit $N \to \infty$. The Matsubara dynamics is derived from the discrete Fourier transform, while the CL-PIMD (2.25) introduced in this paper is obtained from expanding the continuous loop $x(\cdot)$ in the Fourier basis of the Hilbert space $\mathbb{H} = L_p^2([0,\beta];\mathbb{R}^d)$.

2. It is possible to write the CL-PIMD (2.25) directly in terms of the continuous loop $x(\cdot) \in \mathcal{H}$. Suppose $\{\xi^t\}_{t\geqslant 0}$ and $\{\eta^t\}_{t\geqslant 0}$ are evolved by the CL-PIMD (2.25), then we can define

$$x(t,\tau) = \sum_{k=0}^\infty \xi_k^t c_k(\tau), \qquad v(t,\tau) = \sum_{k=0}^\infty \eta_k^t c_k(\tau), \qquad t \geqslant 0, \quad \tau \in [0,\beta], \tag{2.27}$$

where $x(t,\tau)$ and $v(t,\tau)$ represent the position and the velocity of the continuous loop respectively. Assume the damping rates $\gamma_k \equiv \gamma$ for some constant $\gamma > 0$, then multiplying (2.25) by $c_k(\tau)$ and summing the result over k, we formally obtain

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} x(t,\tau) = v(t,\tau), \\ \dfrac{\mathrm{d}}{\mathrm{d}t} v(t,\tau) = -x(t,\tau) - \displaystyle\int_0^\beta K^a(\tau,\tau_0) \nabla V^a(x(t,\tau_0)) \mathrm{d}\tau_0 - \gamma v(t,\tau) + \sqrt{2\gamma} \frac{\mathrm{d}}{\mathrm{d}t} B^a(t,\tau), \end{cases} \tag{2.28}$$

where the kernel function $K^a(\tau,\tau_0)$ and the Brownian motion $B^a(t,\tau)$ are defined by

$$K^a(\tau,\tau_0) = \sum_{k=0}^\infty \frac{c_k(\tau) c_k(\tau_0)}{\omega_k^2 + a^2}, \qquad B^a(t,\tau) = \sum_{k=0}^\infty \frac{c_k(\tau)}{\sqrt{\omega_k^2 + a^2}} B_k(t), \tag{2.29}$$

and $\{B_k(t)\}_{k=0}^\infty$ are independent Brownian motions in $\mathbb{R}^d$. The Langevin dynamics (2.28) has also been previously derived in [22] as a stochastic partial differential equation.

3 Convergence analysis of the CL-PIMD

3.1 Assumptions and convergence results

The CL-PIMD (2.25) is an infinite-dimensional Langevin dynamics of the continuous loop $x(\cdot)$, which is formally defined on the Hilbert space $\mathbb{H} = L_p^2([0,\beta];\mathbb{R}^d)$. We truncate the number of normal modes in the CL-PIMD (2.25) to a finite integer $N \in \mathbb{N}$ to obtain the truncated dynamics,

$$\begin{cases} \dot{\xi}_k = \eta_k, \\ \dot{\eta}_k = -\xi_k - \dfrac{1}{\omega_k^2 + a^2} \displaystyle\int_0^\beta \nabla V^a(x_N(\tau)) c_k(\tau) \mathrm{d}\tau - \gamma_k \eta_k + \sqrt{\dfrac{2\gamma_k}{\omega_k^2 + a^2}} \dot{B}_k, \end{cases} \quad k = 0, 1, \cdots, N-1, \tag{3.1}$$

where the continuous loop $x_N(\cdot) \in \mathbb{H}$ is constructed with the normal mode coordinates $\{\xi_k\}_{k=0}^{N-1}$,

$$x_N(\tau) = \sum_{k=0}^{N-1} \xi_k c_k(\tau). \tag{3.2}$$

The global-in-time existence and well-posedness of the truncated CL-PIMD (3.1) directly follow from the regularity condition on $\nabla V^a(q)$, for example, the global Lipschitz condition. Also, the truncated CL-PIMD (3.1) can be implemented using a suitable numerical discretization scheme, and thus can be applied as sampling method to compute the quantum thermal average.

Now we are mainly intereted in the convergence properties of the truncated CL-PIMD (3.1):

1. How to quantify the difference between the statistical average of the truncated CL-PIMD (3.1) and the quantum thermal average $\langle O(\hat{q}) \rangle_\beta$?
2. How to validate the uniform-in-N geometric ergodicity of the truncated CL-PIMD (3.1)?

The energy function of the truncated Langevin dynamics (3.1) is

$$\mathcal{E}_N(\xi) = \frac{1}{2} \sum_{k=0}^{N-1} (\omega_k^2 + a^2)|\xi_k|^2 + \int_0^\beta V^a\bigg(\sum_{k=0}^{N-1} \xi_k c_k(\tau)\bigg) \mathrm{d}\tau, \quad \xi \in \mathbb{R}^{dN}, \tag{3.3}$$

and the corresponding invariant distribution is the Boltzmann distribution

$$\pi_N(\xi) = \frac{1}{Z_N} \exp(-\mathcal{E}_N(\xi)), \quad Z_N = \int_{\mathbb{R}^{dN}} \exp(-\mathcal{E}_N(\xi)) \mathrm{d}\xi. \tag{3.4}$$

Then the statistical average of the observable function $O(q)$ in $\pi_N(\xi)$ is given by

$$\langle O(\hat{q}) \rangle_{\beta,N} = \int_{\mathbb{R}^{dN}} \bigg[\frac{1}{\beta} \int_0^\beta O\bigg(\sum_{k=0}^{N} \xi_k c_k(\tau)\bigg) \mathrm{d}\tau\bigg] \pi_N(\xi) \mathrm{d}\xi, \tag{3.5}$$

and the difference $|\langle O(\hat{q}) \rangle_{\beta,N} - \langle O(\hat{q}) \rangle_\beta|$ shows the approximation error of the Langevin dynamics (3.1). In the following we consider two typical choices of the damping rates $\{\gamma_k\}_{k=0}^{N-1}$.

- When $\gamma_k = +\infty$ for each normal mode, we obtain the overdamped limit of (3.1), the overdamped Langevin dynamics:

$$\dot{\xi}_k = -\xi_k - \frac{1}{\omega_k^2 + a^2} \int_0^\beta \nabla V^a(x_N(\tau)) c_k(\tau) \mathrm{d}\tau + \sqrt{\frac{2}{\omega_k^2 + a^2}} \dot{B}_k, \quad k = 0, 1, \cdots, N-1. \tag{3.6}$$

- When $\gamma_k = 1$ for each mode, we obtain the underdamped Langevin dynamics with constant damping rates:

$$\begin{cases} \dot{\xi}_k = \eta_k, \\ \dot{\eta}_k = -\xi_k - \dfrac{1}{\omega_k^2 + a^2} \displaystyle\int_0^\beta \nabla V^a(x_N(\tau)) c_k(\tau) \mathrm{d}\tau - \eta_k + \sqrt{\dfrac{2}{\omega_k^2 + a^2}} \dot{B}_k, \end{cases} \quad k = 0, 1, \cdots, N-1. \tag{3.7}$$

Under appropriate confining conditions on the potential function $V(q)$, we can prove that both (3.6) and (3.7) have uniform-in-N geometric ergodicity. Before we conduct a detailed discussion on these results, we enumerate all the assumptions on the potential function $V(q)$ and the observable function $O(q)$ required, and display the convergence results in the overdamped case (3.6) and the underdamped case (3.7) in Table 2 and Table 3.

Assumption For given constant $a > 0$, the potential function $V^a(q) = V(q) - a^2|q|^2/2$ satisfies

(i) $|V^a(0)| \leqslant M_1$, $V^a(q) \geqslant -M_1$ and $|\nabla V^a(q)| \leqslant M_1 + M_1|q|$;

(ii) $V^a(q)$ can be decomposed as $V^c(q) + V^b(q)$, where $V^c(q)$ is convex and $|V^b(q)| \leqslant M_2$;

(iii) $-M_3 I_d \leqslant \nabla^2 V^a(q) \leqslant M_3 I_d$.

The observable function $O(q)$ satisfies

(iv) $\max\{|O(q)|, |\nabla O(q)|, |\nabla^2 O(q)|\} \leqslant M_4$.

Here, M_1, M_2, M_3, M_4 are nonnegative constants, and the functions $V(q)$ and $O(q)$ are assumed to have continuous second order derivatives. Given the scalar function $f \in C^2(\mathbb{R}^d)$, $\nabla f(q) \in \mathbb{R}^d$ is the gradient of $f(q)$, and $\nabla^2 f(q) \in \mathbb{R}^{d\times d}$ is the Hessian matrix of $f(q)$. The symbol $|\cdot|$ means the 2-norm for both vectors in $\mathbb{R}^d$ and matrices in $\mathbb{R}^{d\times d}$, i.e.,

$$|u| = \sqrt{\sum_{p=1}^d |u_p|^2}, \quad u \in \mathbb{R}^d; \quad |B| = \max_{u \in \mathbb{R}^d, u \neq 0} \frac{|Bu|}{|u|}, \quad B \in \mathbb{R}^{d\times d}. \tag{3.8}$$

The partial order relation $A \leqslant B$ for two symmetric matrices $A, B \in \mathbb{R}^{d\times d}$ means the matrix $B - A \in \mathbb{R}^{d\times d}$ is positive semidefinite.

Remark Note that Assumption (iii) implies

$$\nabla V^a(q) - \nabla V^a(0) = \int_0^1 q \cdot \nabla^2 V^a(\theta q) \mathrm{d}\theta \Longrightarrow |\nabla V^a(q)| \leqslant |\nabla V^a(0)| + M_3|q|, \tag{3.9}$$

hence $|\nabla V^a(q)| \leqslant M_1 + M_1|q|$ holds for $M_1 = \max\{|\nabla V^a(0)|, M_3\}$. Also, Assumption (i) implies

$$V^a(q) - V^a(0) = \int_0^1 q \cdot \nabla V^a(\theta q)\mathrm{d}\theta \Longrightarrow |V^a(q)| \leqslant M_1 + |q| \int_0^1 |\nabla V^a(\theta q)|\mathrm{d}\theta. \tag{3.10}$$

And thus we have the following upper bound of $V^a(q)$,

$$V^a(q) \leqslant M_1 + |q| \int_0^1 (M_1 + \theta M_1|q|)\mathrm{d}\theta \leqslant M_1 + M_1|q| + \frac{1}{2}M_1|q|^2 \leqslant \frac{3}{2}M_1 + M_1|q|^2. \tag{3.11}$$

	overdamped Langevin (3.6)	underdamped Langevin (3.7)
assumption	Corollary 3.1: (i)(ii)	Corollary 3.2: (i)(ii)(iii)
ergodicity	$H(\nu_N P_t\|\pi_N) \leqslant C_1 \exp(-2\lambda_1 t)$	$H(\sigma_N P_t\|\mu_N) \leqslant C_2 \exp(-2\lambda_2 t)$

Table 2: The uniform-in-N geometric ergodicity of the overdamped Langevin dynamics (3.6) and the underdamped Langevin dynamics (3.7) in the sense of the relative entropy. ν_N and σ_N are specific initial distributions, and $(P_t)_{t\geqslant 0}$ is the Markov transition semigroup.

The convergence rates λ_1 and λ_2 are given by

$$\lambda_1 = \exp(-4\beta M_2), \qquad \lambda_2 = \frac{a^4}{3M_3^2 + 5a^4}\exp(-4\beta M_2), \tag{3.12}$$

and the constants C_0, C_1 and C_2 are given by

$$C_0 = \frac{d\beta}{2a}\coth\frac{a\beta}{2}, \quad C_1 = \frac{5}{2}\beta M_1 + C_0 M_1 \exp\left(\frac{5}{2}\beta M_1 + C_0 M_1\right), \quad C_2 = C_1 + \frac{4a^4(\beta + C_0)C_0}{d(M_3^2 + a^4)}M_1^2. \tag{3.13}$$

assumption	Theorem 3.4: (i)(iv)
approx. error	$\big\|\langle O(\hat{q})\rangle_{\beta,N} - \langle O(\hat{q})\rangle_\beta\big\| \leqslant \frac{K_1}{N} + \frac{K_2}{\sqrt{N}}$
assumption	Corollary 3.3 (overdamped): (i)(ii)(iv)
sampling error	$\big\|\langle O(\hat{q})\rangle_{\beta,N,t} - \langle O(\hat{q})\rangle_\beta\big\| \leqslant \frac{K_1}{N} + \frac{K_2}{\sqrt{N}} + \sqrt{2C_1}M_4\exp(-\lambda_1 t)$
assumption	Corollary 3.3 (underdamped): (i)(ii)(iii)(iv)
sampling error	$\big\|\langle O(\hat{q})\rangle_{\beta,N,t} - \langle O(\hat{q})\rangle_\beta\big\| \leqslant \frac{K_1}{N} + \frac{K_2}{\sqrt{N}} + \sqrt{2C_2}M_4\exp(-\lambda_2 t)$

Table 3: The approximation error $|\langle O(\hat{q})\rangle_{\beta,N} - \langle O(\hat{q})\rangle_\beta|$, and the sampling error $|\langle O(\hat{q})\rangle_{\beta,N,t} - \langle O(\hat{q})\rangle_\beta|$ for the overdamped Langevin (3.6) and underdamped Langevin (3.7).

The constants K_1 and K_2 are given by

$$K_1 = \frac{d\beta M_4}{8}, \qquad K_2 = 2\sqrt{d(\beta + 2C_0)}\beta M_1 M_4 \exp\left(\frac{5}{2}\beta M_1 + C_0 M_1\right). \tag{3.14}$$

The remaining part of this section devotes to the proof the convergence results in Table 2 and Table 3. In Section 3.2, we prove the uniform-in-N ergodicity of the overdamped Langevin dynamics (3.6) using the log-Sobolev inequalities in the classical Bakry–Émery theory. In Section 3.3, we prove the uniform-in-N ergodicity of the underdamped Langevin dynamics (3.7) using the generalized Γ calculus. In Section 3.4, we quantify the difference between the statistical average $\langle O(\hat{q})\rangle_{\beta,N}$ and the quantum thermal average $\langle O(\hat{q})\rangle_{\beta}$ using the Feynman–Kac formula.

Remark In Table 2, the convergence rate λ_2 is smaller than λ_1. This is because the proof of the ergodicity of the underdamped Langevin dynamics (3.7) using the generalized Γ calculus is actually based on the proof of the overdamped Langevin dynamics (3.6). Also, $\lambda_2 < \lambda_1$ does not imply the overdamped dynamics converges faster than the underdamped dynamics. In fact, it is believed that introducing auxiliary velocity variables will accelerate the convergence of the overdamped Langevin dynamics [37].

3.2 Uniform-in-N ergodicity of overdamped Langevin dynamics

We prove the uniform-in-N geometric ergodicity of the overdamped Langevin dynamics (3.6) in the sense of the relative entropy using the log-Sobolev inequalities in the classical Bakry–Émery theory [26]. Compared to the probabilistic methods, the log-Sobolev inequality is able to produce an explicit convergence rate in high-dimensions.

Theorem 3.1 Under Assumption (ii), let $(P_t)_{t\geqslant 0}$ be the Markov semigroup of the overdamped Langevin dynamics (3.6), then for any positive smooth function $f(\xi)$,

$$\mathrm{Ent}_{\pi_N}(P_t f) \leqslant \exp(-2\lambda_1 t)\,\mathrm{Ent}_{\pi_N}(f), \quad \forall t \geqslant 0, \tag{3.15}$$

where the convergence rate $\lambda_1 = \exp(-4\beta M_2)$.

Here, $\mathrm{Ent}_{\pi_N}(f)$ is the relative entropy of $f(\xi)$ with respect to the distribution $\pi_N(\xi)$,

$$\mathrm{Ent}_{\pi_N}(f) = \int_{\mathbb{R}^{dN}} f \log f \mathrm{d}\pi_N - \int_{\mathbb{R}^{dN}} f \mathrm{d}\pi_N \log \int_{\mathbb{R}^{dN}} f \mathrm{d}\pi_N. \tag{3.16}$$

When $f(\xi)$ is the Radon–Nikodym derivative of the positive distribution $\rho_N(\xi)$ in $\mathbb{R}^{dN}$ with respect to the invariant distribution $\pi_N(\xi)$, i.e.,

$$f(\xi) = \frac{\rho_N(\xi)}{\pi_N(\xi)} > 0, \quad \xi \in \mathbb{R}^{dN}, \tag{3.17}$$

then $\mathrm{Ent}_{\pi_N}(f)$ is exactly the relative entropy $H(\rho_N|\pi_N)$ given by

$$H(\rho_N|\pi_N) = \int_{\mathbb{R}^{dN}} \rho_N(\xi) \log \frac{\rho_N(\xi)}{\pi_N(\xi)} \mathrm{d}\xi. \tag{3.18}$$

Proof The proof is accomplished in several steps. We first prove the uniform-in-N log-Sobolev inequality for the overdamped Langevin dynamics driven by the convex potential function $V^c(q)$, then apply the bounded perturbation property of the log-Sobolev inequalities to obtain the uniform-in-N log-Sobolev inequality for (3.6), which implies the exponential decay of the relative entropy.

1. Derivation of the overdamped Langevin dynamics driven by $V^c(q)$

For notational convenience, introduce

$$\bar{V}^a(\xi) = \int_0^\beta V^a(x_N(\tau))\mathrm{d}\tau, \qquad \bar{V}^c(\xi) = \int_0^\beta V^c(x_N(\tau))\mathrm{d}\tau, \tag{3.19}$$

then the overdamped Langevin dynamics (3.6) can be written as

$$\dot{\xi}_k = -\xi_k - \frac{1}{\omega_k^2 + a^2}\nabla_{\xi_k}\bar{V}^a(\xi) + \sqrt{\frac{2}{\omega_k^2+a^2}}\dot{B}_k, \qquad k = 0, 1, \cdots, N-1. \tag{3.20}$$

To study the geometric ergodicity of (3.20), we first consider its convex part, i.e., the overdamped Langevin dynamics driven by the convex potential function $V^c(q)$:

$$\dot{\xi}_k = -\xi_k - \frac{1}{\omega_k^2 + a^2}\nabla_{\xi_k}\bar{V}^c(\xi) + \sqrt{\frac{2}{\omega_k^2+a^2}}\dot{B}_k, \qquad k = 0, 1, \cdots, N-1. \tag{3.21}$$

The generator of the overdamped Langevin dynamics (3.21) is

$$L^c = -\sum_{k=0}^{N-1}\bigg(\xi_k + \frac{1}{\omega_k^2+a^2}\nabla_{\xi_k}\bar{V}^c(\xi)\bigg)\cdot\nabla_{\xi_k} + \sum_{k=0}^{N-1}\frac{1}{\omega_k^2+a^2}\Delta_{\xi_k}, \tag{3.22}$$

and the invariant distribution of (3.21) is

$$\pi_N^c(\xi) = \frac{1}{Z_N^c}\exp\bigg(-\frac{1}{2}\sum_{k=0}^{N-1}(\omega_k^2+a^2)|\xi_k|^2 - \int_0^\beta V^c(x_N(\tau))\mathrm{d}\tau\bigg), \tag{3.23}$$

and Z_N^c is the normalization constant such that $\int_{\mathbb{R}^{dN}}\pi_N^c(\xi)\mathrm{d}\xi = 1$. We can also verify that $\bar{V}^c(\xi)$ is a convex function of $\{\xi_k\}_{k=0}^{N-1}$. In fact, the second order derivative of $\bar{V}^c(\xi)$ is given by

$$\frac{\partial^2\bar{V}^c(\xi)}{\partial\xi_k\partial\xi_j} = \int_0^\beta \nabla^2 V^c(x_N(\tau))c_k(\tau)c_j(\tau)\mathrm{d}\tau \in \mathbb{R}^{d\times d}, \qquad k, j = 0, 1, \cdots, N-1, \tag{3.24}$$

hence for any real sequence $\{\theta_k\}_{k=0}^{N-1}$ in $\mathbb{R}^d$, the convexity of the potential function $V^c(q)$ implies the Hessian $\nabla^2 V^c(q) \in \mathbb{R}^{d\times d}$ is positive semidefinite and thus

$$\sum_{k,j=0}^{N-1}\theta_k^{\mathrm{T}}\frac{\partial^2\bar{V}^c(\xi)}{\partial\xi_k\partial\xi_j}\theta_j = \int_0^\beta\bigg(\sum_{k=0}^{N-1}\theta_k c_k(\tau)\bigg)^{\mathrm{T}}\nabla^2 V^c(x_N(\tau))\bigg(\sum_{j=0}^{N-1}\theta_j c_j(\tau)\bigg)\mathrm{d}\tau \geqslant 0. \tag{3.25}$$

Then we conclude $\bar{V}^c(\xi)$ is a convex function.

2. Calculation of carré du champ operators $\Gamma_1(f)$ and $\Gamma_2(f)$

Notions of the carré du champ operator $\Gamma_1(f)$ and the iterated operator $\Gamma_2(f)$ are introduced in Definition 1.4.2 and Equation (1.16.1) in [26]. The inequality relations between $\Gamma_1(f)$ and $\Gamma_2(f)$ can easily build the log-Sobolev inequalities and therefore yield the exponential decay of the relative entropy. For notational convenience, define the drift force of the convex part by

$$b_k(\xi) = \xi_k + \frac{1}{\omega_k^2+a^2}\nabla_{\xi_k}\bar{V}^c(\xi), \qquad k = 0, 1, \cdots, N-1, \tag{3.26}$$

then the overdamped Langevin dynamics (3.21) can be simply written as

$$\dot{\xi}_k = -b_k(\xi) + \sqrt{\frac{2}{\omega_k^2 + a^2}}\dot{B}_k, \qquad k = 0, 1, \cdots, N-1. \tag{3.27}$$

Given two smooth functions $f(\xi)$ and $g(\xi)$, the carré du champ operator $\Gamma_1(f, g)$ is given by

$$\begin{aligned}
\Gamma_1(f, g) &= \frac{1}{2}\big(L^c(fg) - fL^c g - gL^c f\big) \\
&= \frac{1}{2}\bigg[-\sum_{k=0}^{N-1} b_k(\xi) \cdot \nabla_{\xi_k}(fg) + \sum_{k=0}^{N-1} \frac{1}{\omega_k^2 + a^2}\Delta_{\xi_k}(fg) \\
&\qquad - f\bigg(-\sum_{k=0}^{N-1} b_k(\xi) \cdot \nabla_{\xi_k} g + \sum_{k=0}^{N-1} \frac{1}{\omega_k^2 + a^2}\Delta_{\xi_k} g\bigg) \\
&\qquad - g\bigg(-\sum_{k=0}^{N-1} b_k(\xi) \cdot \nabla_{\xi_k} f + \sum_{k=0}^{N-1} \frac{1}{\omega_k^2 + a^2}\Delta_{\xi_k} f\bigg)\bigg] \\
&= \sum_{k=0}^{N-1} \frac{1}{\omega_k^2 + a^2}\nabla_{\xi_k} f \cdot \nabla_{\xi_k} g,
\end{aligned}$$

which is the weighted inner product of $\nabla_\xi f, \nabla_\xi g \in \mathbb{R}^{dN}$. Hence $\Gamma_1(f) := \Gamma_1(f, f)$ is given by

$$\Gamma_1(f) = \sum_{k=0}^{N-1} \frac{1}{\omega_k^2 + a^2}|\nabla_{\xi_k} f|^2. \tag{3.28}$$

To compute the iterated operator $\Gamma_2(f, g)$, we decompose $\Gamma_1(f, g)$ into the summation

$$\Gamma_1(f, g) = \sum_{k=0}^{N-1} \Gamma_{1,k}(f, g), \tag{3.29}$$

where each $\Gamma_{1,k}(f, g)$ is calculated as

$$\Gamma_{1,k}(f, g) = \frac{1}{\omega_k^2 + a^2}\nabla_{\xi_k} f \cdot \nabla_{\xi_k} g, \qquad k = 0, 1, \cdots, N-1, \tag{3.30}$$

then the iterated operator $\Gamma_{2,k}(f,g)$ for each $\Gamma_{1,k}(f,g)$ is calculated as

$$
\begin{aligned}
\Gamma_{2,k}(f,g) &= \frac{1}{2}\big(L^c\Gamma_{1,k}(f,g) - \Gamma_{1,k}(f,L^c g) - \Gamma_{1,k}(L^c f,g)\big) \\
&= \frac{1}{2}\Bigg[-\sum_{j=0}^{N-1} b_j(\xi)\cdot\nabla_{\xi_j}\left(\frac{1}{\omega_k^2+a^2}\nabla_{\xi_k} f\cdot\nabla_{\xi_k} g\right) + \sum_{j=0}^{N-1}\frac{1}{\omega_j^2+a^2}\Delta_{\xi_j}\left(\frac{1}{\omega_k^2+a^2}\nabla_{\xi_k} f\cdot\nabla_{\xi_k} g\right) \\
&\qquad - \frac{1}{\omega_k^2+a^2}\nabla_{\xi_k} f\cdot\nabla_{\xi_k}\left(-\sum_{j=0}^{N-1} b_j(\xi)\cdot\nabla_{\xi_j} g + \sum_{j=0}^{N-1}\frac{1}{\omega_k^2+a^2}\Delta_{\xi_j} g\right) \\
&\qquad - \frac{1}{\omega_k^2+a^2}\nabla_{\xi_k} g\cdot\nabla_{\xi_k}\left(-\sum_{j=0}^{N-1} b_j(\xi)\cdot\nabla_{\xi_j} f + \sum_{j=0}^{N-1}\frac{1}{\omega_k^2+a^2}\Delta_{\xi_j} f\right)\Bigg] \\
&= \frac{1}{\omega_k^2+a^2}\Bigg(\sum_{j=0}^{N-1}\frac{1}{\omega_j^2+a^2}\nabla^2_{\xi_j\xi_k} f : \nabla^2_{\xi_j\xi_k} g \\
&\qquad + \frac{1}{2}\sum_{j=0}^{N-1}\nabla_{\xi_k} f\cdot\nabla_{\xi_k} b_j(\xi)\cdot\nabla_{\xi_j} g + \frac{1}{2}\sum_{j=0}^{N-1}\nabla_{\xi_k} g\cdot\nabla_{\xi_k} b_j(\xi)\cdot\nabla_{\xi_j} f\Bigg).
\end{aligned}
$$

Here, the dot product $u\cdot B\cdot v$ for $u\in\mathbb{R}^d$, $B\in\mathbb{R}^{d\times d}$ and $v\in\mathbb{R}^d$ means the matrix product

$$u\cdot B\cdot v = u^{\mathrm{T}}Bv = \sum_{p,q=1}^{d} u_p B_{pq} v_q, \tag{3.31}$$

and the double dot product $A:B$ for $A\in\mathbb{R}^{d\times d}$, $B\in\mathbb{R}^{d\times d}$ means the double inner product

$$A:B = \mathrm{Tr}[A^{\mathrm{T}}B] = \sum_{p,q=1}^{d} A_{pq}B_{pq}. \tag{3.32}$$

Then for the iterated operator $\Gamma_{2,k}(f) := \Gamma_{2,k}(f,f)$, we have

$$\Gamma_{2,k}(f) = \frac{1}{\omega_k^2+a^2}\left(\sum_{j=0}^{N-1}\frac{1}{\omega_j^2+a^2}|\nabla^2_{\xi_j\xi_k} f|^2 + \sum_{j=0}^{N-1}\nabla_{\xi_k} f\cdot\nabla_{\xi_k} b_j(\xi)\cdot\nabla_{\xi_j} f\right). \tag{3.33}$$

Here we note that the matrix $\nabla_{\xi_k} b_j(\xi)\in\mathbb{R}^{d\times d}$ is not necessarily symmetric, hence one cannot exchange the order of $\nabla_{\xi_k} f\in\mathbb{R}^d$ and $\nabla_{\xi_j} f\in\mathbb{R}^d$ in the dot product in (3.33). Finally, the iterated operator $\Gamma_2(f)$ is given by the summation of $\Gamma_{2,k}(f)$, i.e.,

$$\Gamma_2(f) = \sum_{k=0}^{N-1}\Gamma_{2,k}(f). \tag{3.34}$$

Since $\bar{V}^c(\xi)$ is a convex function in ξ, we have the following lower bound estimate of $\Gamma_2(f)$,

$$
\begin{aligned}
\Gamma_2(f) &= \sum_{k=0}^{N-1} \Gamma_{2,k}(f) \\
&= \sum_{k=0}^{N-1} \frac{1}{\omega_k^2 + a^2} \left(\sum_{j=0}^{N-1} \frac{1}{\omega_j^2 + a^2} |\nabla^2_{\xi_j \xi_k} f|^2 + \sum_{j=0}^{N-1} \nabla_{\xi_k} f \cdot \nabla_{\xi_k} b_j(\xi) \cdot \nabla_{\xi_j} f \right) \\
&\geqslant \sum_{k,j=0}^{N-1} \frac{1}{\omega_k^2 + a^2} \nabla_{\xi_k} f \cdot \nabla_{\xi_k} b_j(\xi) \cdot \nabla_{\xi_j} f \\
&= \sum_{k,j=0}^{N-1} \frac{1}{\omega_k^2 + a^2} \nabla_{\xi_k} f \cdot \left(\delta_{kj} + \sum_{j=0}^{N-1} \frac{1}{\omega_j^2 + a^2} \nabla^2_{\xi_k \xi_j} \bar{V}(\xi) \right) \cdot \nabla_{\xi_j} f \\
&= \sum_{k=0}^{N-1} \frac{|\nabla_{\xi_k} f|^2}{\omega_k^2 + a^2} + \sum_{k,j=0}^{N-1} \frac{1}{(\omega_k^2 + a^2)(\omega_j^2 + a^2)} \nabla_{\xi_k} f \cdot \nabla^2_{\xi_k \xi_j} \bar{V}^c(\xi) \cdot \nabla_{\xi_j} f \\
&\geqslant \sum_{k=0}^{N-1} \frac{1}{\omega_k^2 + a^2} |\nabla_{\xi_k} f|^2.
\end{aligned}
$$

Therefore, we obtain the inequality

$$\Gamma_2(f) \geqslant \Gamma_1(f), \quad \text{for any smooth function } f, \tag{3.35}$$

hence the curvature-dimension condition $CD(1, \infty)$ (see Definition 1.16.1 of [26]) holds.

For a general diffusion process $\{X_t\}_{t \geqslant 0}$ with Markov semigroup $(P_t)_{t \geqslant 0}$ and the invariant distribution π, the curvature-dimension condition $CD(\rho, \infty)$ expressed by the functional inequality

$$\Gamma_2(f) \geqslant \rho \Gamma_1(f), \quad \text{for any smooth function } f \tag{3.36}$$

can be derived from a convex potential function, see (1.16.7) and (1.16.8) of [26]. The curvature-dimension condition $CD(\rho, \infty)$ implies the log-Sobolev inequality (see Proposition 5.7.1 of [26])

$$\mathrm{Ent}_\pi(f) \leqslant \frac{1}{2\rho} \int \frac{\Gamma_1(f)}{f} \mathrm{d}\pi, \quad \text{for any positive smooth function } f, \tag{3.37}$$

and thus the exponential decay of the relative entropy (see Theorem 5.2.1 of [26])

$$\mathrm{Ent}_\pi(P_t f) \leqslant e^{-2\rho t} \, \mathrm{Ent}_\pi(f). \tag{3.38}$$

3. Ergodicity of the overdamped Langevin dynamics driven by $V^a(q)$

For a positive smooth function $f(\xi)$, the relative entropy with respect to π_N^c is given by

$$\mathrm{Ent}_{\pi_N^c}(f) = \int_{\mathbb{R}^{dN}} f \log f \mathrm{d}\pi_N^c - \int_{\mathbb{R}^{dN}} f \mathrm{d}\pi_N^c \log \int_{\mathbb{R}^{dN}} f \mathrm{d}\pi_N^c, \tag{3.39}$$

As a consequence of the curvature-dimension condtion $CD(1, \infty)$, the log-Sobolev inequality $LS(1)$ for the distribution $\pi_N^c(\xi)$ holds true (see Proposition 5.7.1 of [26]), i.e.,

$$\mathrm{Ent}_{\pi_N^c}(f) \leqslant \frac{1}{2} \int_{\mathbb{R}^{dN}} \frac{\Gamma_1(f)}{f} \mathrm{d}\pi_N^c = \frac{1}{2} \sum_{k=0}^{N-1} \frac{1}{\omega_k^2 + a^2} \int_{\mathbb{R}^{dN}} \frac{|\nabla_{\xi_k} f|^2}{f} \mathrm{d}\pi_N^c. \tag{3.40}$$

Note that the density functions of the distribution $\pi_N(\xi)$ and $\pi(\xi)$ satisfy

$$\frac{\pi_N^c(\xi)}{\pi_N(\xi)} = \frac{Z_N}{Z_N^c}\exp\left(\int_0^\beta V^b(x_N(\tau))\mathrm{d}\tau\right) \leqslant \exp(\beta M_2) \times \frac{Z_N}{Z_N^c}, \tag{3.41}$$

where we have used $V^b(q) \leqslant M_2$. Now we use $V^b(q) \geqslant -M_2$ to obtain

$$\frac{Z_N}{Z_N^c} = \frac{\displaystyle\int_{\mathbb{R}^{dN}} \exp\left(-\frac{1}{2}\sum_{k=0}^{N-1}(\omega_k^2+a^2)|\xi_k|^2 - \int_0^\beta V^a(x_N(\tau))\mathrm{d}\tau\right)\mathrm{d}\xi}{\displaystyle\int_{\mathbb{R}^{dN}} \exp\left(-\frac{1}{2}\sum_{k=0}^{N-1}(\omega_k^2+a^2)|\xi_k|^2 - \int_0^\beta V^c(x_N(\tau))\mathrm{d}\tau\right)\mathrm{d}\xi} \leqslant \exp(\beta M_2), \tag{3.42}$$

hence from (3.41) and (3.42) we obtain the upper bound

$$\frac{\pi_N^c(\xi)}{\pi_N(\xi)} \leqslant \exp(2\beta M_2). \tag{3.43}$$

Note that we can exchange the postions of π_N and π_N^c (and also Z_N and Z_N^c) in the inequalities (3.41)(3.42)(3.43) and these inequalities remain valid, because we still have $|V^b(q)| \leqslant M_2$. In particular, using the exchanged version of (3.43) we have $\pi_N(\xi)/\pi_N^c(\xi) \leqslant \exp(2\beta M_2)$ and thus

$$\exp(-2\beta M_2) \leqslant \frac{\pi_N^c(\xi)}{\pi_N(\xi)} \leqslant \exp(2\beta M_2), \quad \forall \xi \in \mathbb{R}^{dN}. \tag{3.44}$$

Using the bounded perturbation of the log-Sobolev inequality (Proposition 5.1.6 of [26]), we have

$$\exp(-4\beta M_2)\,\mathrm{Ent}_{\pi_N}(f) \leqslant \frac{1}{2}\sum_{k=0}^{N-1}\frac{1}{\omega_k^2+a^2}\int_{\mathbb{R}^{dN}}\frac{|\nabla_{\xi_k} f|^2}{f}\mathrm{d}\pi_N, \tag{3.45}$$

hence the convergence rate is $\lambda_1 = \exp(-4\beta M_2)$, and the relative entropy has exponential decay,

$$\mathrm{Ent}_{\pi_N}(P_t f) \leqslant \exp(-2\lambda_1 t)\,\mathrm{Ent}_{\pi_N}(f), \quad \text{for any positive smooth function } f(\xi). \tag{3.46}$$

□

Using Theorem 3.1, we can prove that the relative entropy $H(\nu_N P_t|\pi_N)$ decays exponentially in time, and the convergence rate λ_1 does not depend on the choice of the initial distribution ν_N. For simplicity, we set ν_N to be the Gaussian distribution in the following corollary.

Corollary 3.1 Under Assumptions (i)(ii), if the initial distribution $\nu_N(\xi)$ is chosen as

$$\xi_k \sim \mathcal{N}\left(0, \frac{I_d}{\omega_k^2+a^2}\right), \quad k = 0, 1, \cdots, N-1, \tag{3.47}$$

then the relative entropy $H(\nu_N P_t|\pi_N)$ has exponential decay in time,

$$H(\nu_N P_t|\pi_N) \leqslant C_1 \exp(-2\lambda_1 t), \quad \forall t \geqslant 0, \tag{3.48}$$

where the convergence rate $\lambda_1 = \exp(-4\beta M_2)$ and the constants

$$C_0 = \frac{d\beta}{2a}\coth\frac{a\beta}{2}, \quad C_1 = \frac{5}{2}\beta M_1 + C_0 M_1 \exp\left(\frac{5}{2}\beta M_1 + C_0 M_1\right). \tag{3.49}$$

The proof is left in Appendix A. Note the the choice of ν_N simplifies the derivation of the constant C_1, which explicitly shows that C_1 is independent of the number of normal models N.

3.3 Uniform-in-N ergodicity of underdamped Langevin dynamics

We prove the uniform-in-N geometric ergodicity of the underdamped Langevin dynamics (3.7) in the sense of the relative entropy, and the main technique is the generalized Γ calculus developed in [27, 28]. The generalized Γ calculus is an extension of the classical Bakry–Émery theory and can be applied on stochastic processes with degenerate diffusions. The generalized Γ calculus is largely inspired from the hypocoercivity theory [38] of Villani, and is able to produce an explicit convergence rate in the relative entropy rather than H^1 or L^2. For convenience, we present a brief review of the theory of the generalized Γ calculus in Appendix C. Also note that the proof of the ergodicity of the underdamped case (3.7) directly relies on the results in Section 3.2.

The Hamiltonian function of the underdamped Langevin dynamics (3.7) is

$$\mathcal{H}_N(\xi,\eta) = \frac{1}{2}\sum_{k=0}^{N-1}(\omega_k^2+a^2)(|\xi_k|^2+|\eta_k|^2) + \int_0^\beta V^a\bigg(\sum_{k=0}^{N-1}\xi_k c_k(\tau)\bigg)\mathrm{d}\tau, \tag{3.50}$$

and the corresponding Boltzmann distribution is

$$\mu_N(\xi,\eta) = \frac{1}{Z_N'}\exp(-\mathcal{H}_N(\xi,\eta)), \qquad Z_N' = \int_{\mathbb{R}^{2dN}}\exp(-\mathcal{H}_N(\xi,\eta))\mathrm{d}\xi\mathrm{d}\eta. \tag{3.51}$$

Now we prove the uniform-in-N ergodicity using the generalized Γ calculus [27, 28].

Theorem 3.2 Under Assumptions (ii)(iii), let $(P_t)_{t\geqslant 0}$ be the Markov semigroup of the underdamped Langevin dynamics (3.7). For any positive smooth function $f(\xi,\eta)$, define

$$W(f) = \bigg(\frac{M_3^2}{a^4}+1\bigg)\mathrm{Ent}_{\mu_N}(f) + \sum_{k=0}^{N-1}\frac{1}{\omega_k^2+a^2}\int_{\mathbb{R}^{2dN}}\frac{|\nabla_{\eta_k}f-\nabla_{\xi_k}f|^2+|\nabla_{\eta_k}f|^2}{f}\mathrm{d}\mu_N, \tag{3.52}$$

then

$$W(P_tf) \leqslant \exp(-2\lambda_2 t)W(f), \qquad \forall t\geqslant 0, \tag{3.53}$$

where the convergence rate $\lambda_2 = \frac{a^4}{3M_3^2+5a^4}\exp(-4\beta M_2)$.

Proof The proof is accomplished in several steps. We first establish the uniform-in-N log-Sobolev inequality for the distribution $\mu_N(\xi,\eta)$, then introduce the functions

$$\Phi_1(f) = f\log f, \qquad \Phi_2(f) = \sum_{k=0}^{N-1}\frac{1}{\omega_k^2+a^2}\frac{|\nabla_{\eta_k}f-\nabla_{\xi_k}f|^2+|\nabla_{\eta_k}f|^2}{f}. \tag{3.54}$$

By validating the generalized curvature-dimension condition

$$\Gamma_{\Phi_2}(f) - \frac{1}{2}\Phi_2(f) + \bigg(\frac{M_3^2}{a^4}+1\bigg)\Gamma_{\Phi_1}(f) \geqslant 0, \tag{3.55}$$

we can prove the exponential decay of the quantity $W(P_tf)$.

1. Log-Sobolev inequality for $\mu_N(\xi,\eta)$

Recall that the log-Sobolev inequality for π_N in (3.45) holds true:

$$\lambda_1 \,\mathrm{Ent}_{\pi_N}(f) \leqslant \frac{1}{2}\sum_{k=0}^{N-1}\frac{1}{\omega_k^2+a^2}\int_{\mathbb{R}^{dN}}\frac{|\nabla_{\xi_k} f|^2}{f}\mathrm{d}\pi_N, \quad \text{for any positive function } f(\xi), \tag{3.56}$$

where the convergence rate $\lambda_1 = \exp(-4\beta M_2)$. For the velocity variables $\{\eta_k\}_{k=0}^{N-1}$, consider the following Ornstein–Uhlenbeck process

$$\dot{\eta}_k = -\eta_k + \sqrt{\frac{2}{\omega_k^2+a^2}}\dot{B}_k', \quad k = 0, 1, \cdots, N-1, \tag{3.57}$$

where $\{B_k'\}_{k=0}^{N-1}$ is another set of independent Brownian motions in $\mathbb{R}^d$. The invariant distribution of the Ornstein–Uhlenbeck process (3.57) is exactly the Gaussian distribution $\nu_N(\eta)$,

$$\nu_N(\eta) = \prod_{k=0}^{N-1}\left(\frac{\omega_k^2+a^2}{2\pi}\right)^{\frac{d}{2}}\exp\bigg(-\frac{1}{2}(\omega_k^2+a^2)|\eta_k|^2\bigg), \quad \eta \in \mathbb{R}^{dN}. \tag{3.58}$$

Note that the Ornstein–Uhlenbeck process (3.57) can be viewed as the overdamped Langevin dynamics (3.6) with the potential function $V^a(q) \equiv 0$, hence using again Theorem 3.1 with the constant $M_2 = 0$, the log-Sobolev inequality for ν_N holds true,

$$\mathrm{Ent}_{\nu_N}(f) \leqslant \frac{1}{2}\sum_{k=0}^{N-1}\frac{1}{\omega_k^2+a^2}\int_{\mathbb{R}^{dN}}\frac{|\nabla_{\eta_k} f|^2}{f}\mathrm{d}\nu_N, \quad \text{for any positive function } f(\eta). \tag{3.59}$$

Using the tensorization of log-Sobolev inequalities (Proposition 5.2.7 of [26]), (3.56) and (3.59) yield

$$\lambda_1 \mathrm{Ent}_{\pi_N}(f) \leqslant \frac{1}{2}\sum_{k=0}^{N-1}\frac{1}{\omega_k^2+a^2}\int_{\mathbb{R}^{2dN}}\frac{|\nabla_{\xi_k} f|^2+|\nabla_{\eta_k} f|^2}{f}\mathrm{d}\mu_N, \text{ for any positive function } f(\xi,\eta). \tag{3.60}$$

Here, we construct the tensorized log-Sobolev inequality on the product space $\mathbb{R}^{dN}\times\mathbb{R}^{dN}$, and the convergence rate is determined by the smaller one of $\lambda_1 = \exp(-4\beta M_2)$ and 1, which is λ_1 itself. Note that the log-Sobolev inequality (3.60) does not imply the ergodicity of the underdamped Langevin dynamics (3.7) directly, because (3.7) has degenerate diffusion in the η variable.

2. Calculation of commutators $[L, \nabla_{\xi_k}]$ and $[L, \nabla_{\eta_k}]$

We still use the notation

$$\bar{V}^a(\xi) = \int_0^\beta V^a\bigg(\sum_{k=0}^{N-1}\xi_k c_k(\tau)\bigg)\mathrm{d}\tau, \tag{3.61}$$

and the generator of the underdamped Langevin dynamics (3.7) is given by

$$L = \sum_{k=0}^{N-1}\eta_k\cdot\nabla_{\xi_k} - \sum_{k=0}^{N-1}\bigg(\xi_k+\eta_k+\frac{1}{\omega_k^2+a^2}\nabla_{\xi_k}\bar{V}^a(\xi)\bigg)\cdot\nabla_{\eta_k} + \sum_{k=0}^{N-1}\frac{\Delta_{\xi_k}}{\omega_k^2+a^2}. \tag{3.62}$$

Next we compute the commutators $[L, \nabla_{\xi_k}]$ and $[L, \nabla_{\eta_k}]$. For any $k = 0, 1, \cdots, N-1$,

$$
\begin{aligned}
L\nabla_{\xi_k} &= \sum_{j=0}^{N-1} \eta_j \cdot \nabla^2_{\xi_j\xi_k} - \sum_{j=0}^{N-1} \left(\xi_j + \eta_j + \frac{1}{\omega_j^2 + a^2} \nabla_{\xi_j} \bar{V}^a(\xi) \right) \cdot \nabla_{\xi_k} \nabla_{\eta_j} + \sum_{j=0}^{N-1} \frac{\Delta_{\xi_j} \nabla_{\xi_k}}{\omega_j^2 + a^2}, \\
\nabla_{\xi_k} L &= \sum_{j=0}^{N-1} \eta_j \cdot \nabla^2_{\xi_j\xi_k} - \nabla_{\eta_k} - \sum_{j=0}^{N-1} \frac{1}{\omega_j^2 + a^2} \nabla^2_{\xi_j\xi_k} \bar{V}^a(\xi) \cdot \nabla_{\eta_j} \\
&\quad - \sum_{j=0}^{N-1} \left(\xi_j + \eta_j + \frac{1}{\omega_j^2 + a^2} \nabla_{\xi_j} \bar{V}^a(\xi) \right) \cdot \nabla_{\xi_k} \nabla_{\eta_j} + \sum_{j=0}^{N-1} \frac{\Delta_{\xi_j} \nabla_{\xi_k}}{\omega_j^2 + a^2},
\end{aligned}
$$

hence the commutator $[L, \nabla_{\xi_k}] = L\nabla_{\xi_k} - \nabla_{\xi_k} L$ is given by

$$
[L, \nabla_{\xi_k}] = \nabla_{\eta_k} + \sum_{j=0}^{N-1} \frac{1}{\omega_j^2 + a^2} \nabla^2_{\xi_j\xi_k} \bar{V}^a(\xi) \cdot \nabla_{\eta_j}. \tag{3.63}
$$

Similarly, we can calculate

$$
\begin{aligned}
[L, \nabla_{\eta_k}] &= L\nabla_{\eta_k} - \nabla_{\eta_k} L \\
&= \sum_{j=0}^{N-1} \eta_j \nabla_{\xi_j} \nabla_{\eta_k} - \sum_{k=0}^{N-1} (\xi_j + \eta_j) \cdot \nabla^2_{\eta_j\eta_k} + \sum_{j=0}^{N-1} \frac{\Delta_{\xi_j} \nabla_{\eta_k}}{\omega_j^2 + a^2} \\
&\quad - \nabla_{\xi_k} - \sum_{j=0}^{N-1} \eta_j \nabla_{\xi_j} \nabla_{\eta_k} + \nabla_{\eta_k} + \sum_{j=0}^{N-1} (\xi_j + \eta_j) \nabla^2_{\eta_j\eta_k} - \sum_{j=0}^{N-1} \frac{\Delta_{\xi_j} \nabla_{\eta_k}}{\omega_j^2 + a^2} \\
&= \nabla_{\eta_k} - \nabla_{\xi_k}.
\end{aligned} \tag{3.64}
$$

3. Calculation of generalized Γ operators for $\Phi_1(f)$ and $\Phi_2(f)$

Inspired from Example 3 of [28], we consider the following two functions

$$
\Phi_1(f) = f \log f, \qquad \Phi_2(f) = \sum_{k=0}^{N-1} \frac{1}{\omega_k^2 + a^2} \frac{|\nabla_{\eta_k} f - \nabla_{\xi_k} f|^2 + |\nabla_{\eta_k} f|^2}{f}. \tag{3.65}
$$

See also (C.41) of Appendix C. Here, $\Phi_2(f)$ can be viewed as a twisted form of

$$
\sum_{k=0}^{N-1} \frac{1}{\omega_k^2 + a^2} \frac{|\nabla_{\xi_k} f|^2 + |\nabla_{\eta_k} f|^2}{f}, \tag{3.66}
$$

which appears in the RHS of the log-Sobolev inequality (3.60). The weights $1/(\omega_k^2 + a^2)$ are chosen according to the diffusion coefficients of the underdamped Langevin dynamics (3.7).

Next we calculate the Γ operators $\Gamma_{\Phi_1}(f)$ and $\Gamma_{\Phi_2}(f)$. From Example C.2 of Appendix C,

$$
\Gamma_{\Phi_1}(f) = \frac{1}{2} \sum_{k=0}^{N-1} \frac{1}{\omega_k^2 + a^2} \frac{|\nabla_{\eta_k} f|^2}{f}. \tag{3.67}
$$

To compute $\Gamma_{\Phi_2}(f)$, we write $\Phi_2(f)$ as the summation

$$\Phi_2(f) = \sum_{k=0}^{N-1} \frac{1}{\omega_k^2 + a^2} \Phi_{2,k}(f), \qquad \Phi_{2,k}(f) = \frac{|\nabla_{\eta_k} f - \nabla_{\xi_k} f|^2 + |\nabla_{\eta_k} f|^2}{f}. \tag{3.68}$$

For notational convenience, for a given positive smooth function $f(\xi,\eta)$, we simply write

$$f_{\xi_k} = \nabla_{\xi_k} f, \qquad f_{\eta_k} = \nabla_{\eta_k} f, \qquad k = 0, 1, \cdots, N-1. \tag{3.69}$$

According to Example C.4 in Appendix C, we have

$$\begin{aligned} f \cdot \Gamma_{\Phi_{2,k}}(f) &\geqslant (\nabla_{\eta_k} f - \nabla_{\xi_k} f) \cdot [L, \nabla_{\eta_k} - \nabla_{\xi_k}] f + \nabla_{\eta_k} f \cdot [L, \nabla_{\eta_k}] f \\ &\geqslant (f_{\eta_k} - f_{\xi_k}) \cdot \left(- f_{\xi_k} - \sum_{j=0}^{N-1} \frac{1}{\omega_j^2 + a^2} \nabla^2_{\xi_j \xi_k} \bar{V}^a(\xi) \cdot f_{\eta_j} \right) + f_{\eta_k} \cdot (f_{\eta_k} - f_{\xi_k}) \\ &= |f_{\eta_k} - f_{\xi_k}|^2 - (f_{\eta_k} - f_{\xi_k}) \cdot \sum_{j=0}^{N-1} \frac{1}{\omega_j^2 + a^2} \nabla^2_{\xi_j \xi_k} \bar{V}^a(\xi) \cdot f_{\eta_j}. \end{aligned} \tag{3.70}$$

Hence the summation of (3.70) over $k = 0, 1, \cdots, N-1$ gives

$$f \cdot \Gamma_{\Phi_2}(f) \geqslant \sum_{k=0}^{N-1} \frac{1}{\omega_k^2 + a^2} |f_{\eta_k} - f_{\xi_k}|^2 - \sum_{k,j=0}^{N-1} \frac{f_{\eta_k} - f_{\xi_k}}{\omega_k^2 + a^2} \cdot \nabla^2_{\xi_j \xi_k} \bar{V}^a(\xi) \cdot \frac{f_{\eta_j}}{\omega_j^2 + a^2}. \tag{3.71}$$

Define the vectors $X, Y \in \mathbb{R}^{dN}$ to be

$$X = \left\{ \frac{f_{\eta_k} - f_{\xi_k}}{\sqrt{\omega_k^2 + a^2}} \right\}_{k=0}^{N-1} \in \mathbb{R}^{dN}, \qquad Y = \left\{ \frac{f_{\eta_k}}{\sqrt{\omega_k^2 + a^2}} \right\}_{k=0}^{N-1} \in \mathbb{R}^{dN}, \tag{3.72}$$

then $\Gamma_{\Phi_2}(f)$ can be simply written as

$$\Gamma_{\Phi_2}(f) \geqslant \frac{|X|^2 - X^{\mathrm{T}} \Sigma Y}{f}, \tag{3.73}$$

where the matrix $\Sigma \in \mathbb{R}^{dN} \times \mathbb{R}^{dN}$ is given by

$$\Sigma_{kj} = \frac{1}{\sqrt{(\omega_k^2 + a^2)(\omega_j^2 + a^2)}} \nabla^2_{\xi_k \xi_j} \bar{V}^a(\xi) \in \mathbb{R}^{d \times d}. \tag{3.74}$$

Next we prove that Assumption (iii) implies

$$-\frac{M_3}{a^2} I_{dN} \leqslant \Sigma \leqslant \frac{M_3}{a^2} I_{dN}. \tag{3.75}$$

In fact, for any real sequence $\{\theta_k\}_{k=0}^{N-1}$ in $\mathbb{R}^d$, we have

$$\begin{aligned} \sum_{k,j=0}^{N-1} \theta_k^{\mathrm{T}} \Sigma \theta_j &= \sum_{k,j=0}^{N-1} \int_0^\beta \theta_k^{\mathrm{T}} \nabla^2 V^a(x_N(\tau)) \theta_j \frac{c_k(\tau)}{\sqrt{\omega_k^2 + a^2}} \frac{c_j(\tau)}{\sqrt{\omega_j^2 + a^2}} \mathrm{d}\tau \\ &= \int_0^\beta \left(\sum_{k=0}^{N-1} \frac{\theta_k c_k(\tau)}{\sqrt{\omega_k^2 + a^2}} \right)^{\mathrm{T}} \nabla^2 V^a(x_N(\tau)) \left(\sum_{j=0}^{N-1} \frac{\theta_j c_j(\tau)}{\sqrt{\omega_j^2 + a^2}} \right) \mathrm{d}\tau, \end{aligned}$$

hence using $-M_3 I_d \leqslant \nabla^2 V^a(q) \leqslant M_3 I_d$, we have

$$\sum_{k,j=0}^{N-1} \theta_k^{\mathrm{T}} \Sigma \theta_j \leqslant M_3 \int_0^\beta \bigg| \sum_{k=0}^{N-1} \frac{\theta_k c_k(\tau)}{\sqrt{\omega_k^2 + a^2}} \bigg|^2 \mathrm{d}\tau = M_3 \sum_{k=0}^{N-1} \frac{|\theta_k|^2}{\omega_k^2 + a^2} \leqslant \frac{M_3}{a^2} \sum_{k=0}^{N-1} |\theta_k|^2, \tag{3.76}$$

and similarly

$$\sum_{k,j=0}^{N-1} \theta_k^{\mathrm{T}} \nabla^2 \bar{V}^a(\xi) \theta_j \geqslant -\frac{M_3}{a^2} \sum_{k=0}^{N-1} |\theta_k|^2. \tag{3.77}$$

Hence from (3.73) we obtain

$$\Gamma_{\Phi_2}(f) \geqslant \frac{|X|^2 - \frac{M_3}{a^2}|X||Y|}{f}. \tag{3.78}$$

In conclusion, the functions $\Phi_1(f)$ and $\Phi_2(f)$ satisfy

$$\Phi_1(f) = f \log f, \qquad \Phi_2(f) = \frac{|X|^2 + |Y|^2}{f}. \tag{3.79}$$

$$\Gamma_{\Phi_1}(f) = \frac{|Y|^2}{2f}, \qquad \Gamma_{\Phi_2}(f) \geqslant \frac{|X|^2 - \frac{M_3}{a^2}|X||Y|}{f}, \tag{3.80}$$

where the vectors $X, Y \in \mathbb{R}^{dN}$ are given in (3.72).

4. Construction of functional inequalities

Let us summarize the functional inequalities so far. The log-Sobolev inequality (3.60) implies

$$\lambda_1 \operatorname{Ent}_{\mu_N}(f) \leqslant \frac{1}{2} \int_{\mathbb{R}^{2dN}} \frac{|X+Y|^2 + |Y|^2}{f} \mathrm{d}\mu_N \leqslant \frac{3}{2} \int_{\mathbb{R}^{2dN}} \frac{|X|^2 + |Y|^2}{f} \mathrm{d}\mu_N, \tag{3.81}$$

hence

$$\frac{2\lambda_1}{3} \bigg(\int_{\mathbb{R}^{2dN}} \Phi_1(f) \mathrm{d}\mu_N - \Phi_1\bigg(\int_{\mathbb{R}^{2dN}} f \mathrm{d}\mu_N \bigg) \bigg) \leqslant \int_{\mathbb{R}^{2dN}} \Phi_2(f) \mathrm{d}\mu_N. \tag{3.82}$$

Also,

$$\Gamma_{\Phi_2}(f) - \frac{1}{2}\Phi_2(f) \geqslant \frac{|X|^2 - 2\frac{M_3}{a^2}|X||Y| - |Y|^2}{2f}, \tag{3.83}$$

and thus

$$\Gamma_{\Phi_2}(f) - \frac{1}{2}\Phi_2(f) + \bigg(\frac{M_3^2}{a^4} + 1 \bigg) \Gamma_{\Phi_1}(f) \geqslant \frac{(|X| - \frac{M_3}{a^2}|Y|)^2}{2f} \geqslant 0. \tag{3.84}$$

Recall that $(P_t)_{t \geqslant 0}$ is the Markov semigroup of the underdamped Langevin dynamics (3.7). Using Theorem C.1 in Appendix C, we define the quantity

$$\begin{aligned} W(f) &= \bigg(\frac{M_3^2}{a^4} + 1 \bigg) \bigg(\int_{\mathbb{R}^{2dN}} \Phi_1(f) \mathrm{d}\mu_N - \Phi_1\bigg(\int_{\mathbb{R}^{2dN}} f \mathrm{d}\mu_N \bigg) \bigg) + \int_{\mathbb{R}^{2dN}} \Phi_2(f) \mathrm{d}\mu_N \\ &= \bigg(\frac{M_3^2}{a^4} + 1 \bigg) \operatorname{Ent}_{\mu_N}(f) + \sum_{k=0}^{N-1} \frac{1}{\omega_k^2 + a^2} \int_{\mathbb{R}^{2dN}} \frac{|\nabla_{\eta_k} f - \nabla_{\xi_k} f|^2 + |\nabla_{\eta_k} f|^2}{f} \mathrm{d}\mu_N. \end{aligned}$$

Then using the functional inequalities (3.82) and (3.84), we have

$$W(P_t f) \leqslant \exp\bigg(-\frac{t}{1+\frac{3(M_3^2/a^4+1)}{2\lambda_1}}\bigg)W(f), \quad \forall t \geqslant 0. \tag{3.85}$$

Note that the convergence rate satisfies

$$\frac{1}{1+\frac{3(M_3^2/a^4+1)}{2\lambda_1}} = \frac{1}{1+\frac{3(M_3^2/a^4+1)}{2}\exp(4\beta M_2)} \geqslant \frac{2a^4}{3M_3^2+5a^4}\exp(-4\beta M_2), \tag{3.86}$$

hence for $\lambda_2 = a^4 \exp(-4\beta M_2)/(3M_3^2+5a^4)$ we obtain

$$W(P_t f) \leqslant \exp(-2\lambda_2 t)W(f), \quad \forall t \geqslant 0, \tag{3.87}$$

which completes the proof. □

Using Theorem 3.2, we can prove that the quantity $W(P_t f)$ decays exponentially in time, and the convergence rate λ_2 does not depend on the choice of the initial distribution σ_N. For simplicity, we set the initial distribution σ_N to be the Gaussian distribution in the following corollary.

Corollary 3.2 Under Assumptions (i)(ii)(iii), if the initial distribution $\sigma_N(\xi,\eta)$ is chosen as

$$\xi_k, \eta_k \sim \mathcal{N}\bigg(0, \frac{I_d}{\omega_k^2+a^2}\bigg), \quad k = 0, 1, \cdots, N-1, \tag{3.88}$$

then the relative entropy $H(\sigma_N P_t|\mu_N)$ has exponential decay in time,

$$H(\sigma_N P_t|\mu_N) \leqslant C_2 \exp(-2\lambda_2 t), \quad \forall t \geqslant 0, \tag{3.89}$$

where the convergence rate $\lambda_2 = \frac{a^4}{3M_3^2+5a^4}\exp(-4\beta M_2)$, and the constant

$$C_0 = \frac{d\beta}{2a}\coth\frac{a\beta}{2}, \quad C_2 = \frac{5}{2}\beta M_1 + C_0 M_1 \exp\bigg(\frac{5}{2}\beta M_1 + C_0 M_1\bigg) + \frac{4a^4(\beta+C_0)C_0}{d(M_3^2+a^4)}M_1. \tag{3.90}$$

The proof is left in Appendix A. Note the the choice of σ_N simplifies the derivation of the constant C_2, which explicitly shows that C_2 is independent of the number of normal models N.

3.4 Error in computing the quantum thermal average $\langle O(\hat{q})\rangle_\beta$

We quantify the difference between $\langle O(\hat{q})\rangle_{\beta,N}$ and $\langle O(\hat{q})\rangle_\beta$ using the Feynman–Kac formula [24].

Lemma 3.1 Under Assumptions (i)(iv), for any positive integers $N, M \in \mathbb{N}$ with $M > N$,

$$|\langle O(\hat{q})\rangle_{\beta,N} - \langle O(\hat{q})\rangle_{\beta,M}| \leqslant \frac{d\beta M_4}{8N} + 2\sqrt{d(\beta+2C_0)}\beta M_1 M_4 \exp\bigg(\frac{5}{2}\beta M_1 + C_0 M_1\bigg)\frac{1}{\sqrt{N}}, \tag{3.91}$$

where the constant $C_0 = \frac{d\beta}{2a}\coth\frac{a\beta}{2}$.

The proof of Lemma 3.1 is straightforward, and the details are left in Appendix B. Lemma 3.1 shows that $\{\langle O(\hat{q})\rangle_{\beta,N}\}_{N=1}^{\infty}$ is a Cauchy sequence, and thus has a limit as $N\to\infty$. Nevertheless, it requires additional arguments to show the limit coincides with the quantum average $\langle O(\hat{q})\rangle_{\beta}$.

On the one hand, the Hilbert space $\mathbb{H}=L_p^2([0,\beta];\mathbb{R}^d)$ defined in (2.11) can be equivalently written in the normal mode coordinates

$$(\xi,\eta)_{\mathbb{H}}=\sum_{k=0}^{\infty}\xi_k\eta_k,\qquad\forall\xi,\eta\in\mathbb{H}.\tag{3.92}$$

Introduce the distribution ν of the real sequence $\xi=\{\xi_k\}_{k=0}^{\infty}$ as

$$\xi_k\sim\mathcal{N}\bigg(0,\frac{I_d}{\omega_k^2+a^2}\bigg),\qquad k=0,1,2,\cdots,\tag{3.93}$$

then in the distribution ν, the continuous loop

$$x_N(\tau)=\sum_{k=0}^{N-1}\xi_kc_k(\tau)\in\mathbb{H},\qquad\tau\in[0,\beta]\tag{3.94}$$

can be viewed as a random variable in $\mathbb{H}$ whose distribution law is the pushforward of ν onto $\mathbb{H}$ through the relation (3.94). Next, the definition of the invariant distribution $\pi_N(\xi)$ in (3.4) implies the statistical average $\langle O(\hat{q})\rangle_{\beta,N}$ can be equivalently written as

$$\langle O(\hat{q})\rangle_{\beta,N}=\frac{\mathbb{E}_{\nu}\bigg[\exp\bigg(-\displaystyle\int_0^{\beta}V^a(x_N(\tau))\mathrm{d}\tau\bigg)\times\bigg(\frac{1}{\beta}\displaystyle\int_0^{\beta}O(x_N(\tau))\mathrm{d}\tau\bigg)\bigg]}{\mathbb{E}_{\nu}\bigg[\exp\bigg(-\displaystyle\int_0^{\beta}V^a(x_N(\tau))\mathrm{d}\tau\bigg)\bigg]}.\tag{3.95}$$

On the other hand, for any positive integers $N,M\in\mathbb{N}$ with $M>N$, it is easy to see

$$\mathbb{E}_{\nu}|x_N(\tau)-x_M(\tau)|^2=\sum_{k=N}^{M-1}\mathbb{E}|\xi_k|^2=\sum_{k=N}^{M-1}\frac{d}{\omega_k^2+a^2},\tag{3.96}$$

hence $\{x_N(\cdot)\}_{N=1}^{\infty}$ is a Cauchy sequence in $L^2(\nu)$, and we can identify the limit $x(\tau)$ in $L^2(\nu)$ as

$$x(\tau)=\sum_{k=0}^{\infty}\xi_kc_k(\tau),\qquad\tau\in[0,\beta].\tag{3.97}$$

Moreover, for any $\tau_1,\tau_2\in[0,\beta]$, we have

$$\begin{aligned}\mathbb{E}_{\nu}\big|x(\tau_1)-x(\tau_2)\big|^2&=\mathbb{E}_{\nu}\bigg|\sum_{k=0}^{\infty}\xi_kc_k(\tau_1)-\sum_{k=0}^{\infty}\xi_kc_k(\tau_2)\bigg|^2\\&=\sum_{k=0}^{\infty}|c_k(\tau_1)-c_k(\tau_2)|^2\mathbb{E}|\xi_k|^2=\sum_{k=0}^{\infty}\frac{d}{\omega_k^2+a^2}|c_k(\tau_1)-c_k(\tau_2)|^2\\&=\frac{2d}{\beta}\sum_{k=1}^{\infty}\frac{\big(\sin\frac{2k\pi\tau_1}{\beta}-\sin\frac{2k\pi\tau_2}{\beta}\big)^2}{(\frac{2k\pi}{\beta})^2+a^2}+\frac{2d}{\beta}\sum_{k=1}^{\infty}\frac{\big(\cos\frac{2k\pi\tau_1}{\beta}-\cos\frac{2k\pi\tau_2}{\beta}\big)^2}{(\frac{2k\pi}{\beta})^2+a^2}\\&=\frac{8d}{\beta}\sum_{k=1}^{\infty}\frac{\sin^2\frac{k\pi(\tau_1-\tau_2)}{\beta}}{(\frac{2k\pi}{\beta})^2+a^2}\leqslant\frac{8d}{\beta}\sum_{k=1}^{\infty}\frac{k\pi|\tau_1-\tau_2|}{\beta}\frac{1}{(\frac{2k\pi}{\beta})^2+a^2}\leqslant C|\tau_1-\tau_2|,\end{aligned}$$

hence we conclude that there exists a constant C independent of τ_1, τ_2 such that

$$\mathbb{E}_\nu |x(\tau_1) - x(\tau_2)|^2 \leqslant C|\tau_1 - \tau_2|. \tag{3.98}$$

Using the Kolmogorov continuity theorem, $x(\tau)$ can be viewed as a locally α-Hölder continuous function for any $0 < \alpha < 1/2$. In total, $x(\tau)$ is a random continuous path in the Hilbert space $\mathbb{H}$ defined by the $L^2(\nu)$ limit of the infinite summation in (3.97). Therefore, we expect the limit $\lim_{N\to\infty} \langle O(\hat{q})\rangle_{\beta,N}$ can be represented by the expectation of the random continuous loop $x(\tau)$.

Lemma 3.2 Under Assumptions (i)(iv), the limit of the statistical average $\langle O(\hat{q})\rangle_{\beta,N}$ is

$$\lim_{N\to\infty} \langle O(\hat{q})\rangle_{\beta,N} = \frac{\mathbb{E}_\nu\bigg[\exp\bigg(-\displaystyle\int_0^\beta V^a(x(\tau))\mathrm{d}\tau\bigg) \times \bigg(\frac{1}{\beta}\displaystyle\int_0^\beta O(x(\tau))\mathrm{d}\tau\bigg)\bigg]}{\mathbb{E}_\nu\bigg[\exp\bigg(-\displaystyle\int_0^\beta V^a(x(\tau))\mathrm{d}\tau\bigg)\bigg]}. \tag{3.99}$$

The proof of Lemma 3.2 mainly employs (3.95) and the $L^2(\nu)$ convergence of $x_N(\tau) \in \mathbb{H}$ towards $x(\tau) \in \mathbb{H}$, and the details of the proof are left in Appendix B.

Finally, the quantum thermal average $\langle O(\hat{q})\rangle_\beta$ is related to the RHS of (3.99) in the following result, which can be viewed as an extension of the Feynman–Kac formula.

Theorem 3.3 Under Assumptions (i)(iv), the quantum thermal average $\langle O(\hat{q})\rangle_\beta$ is represented as

$$\langle O(\hat{q})\rangle_\beta = \frac{\mathrm{Tr}[e^{-\beta\hat{H}}O(\hat{q})]}{\mathrm{Tr}[e^{-\beta\hat{H}}]} = \frac{\mathbb{E}_\nu\bigg[\exp\bigg(-\displaystyle\int_0^\beta V^a(x(\tau))\mathrm{d}\tau\bigg) \times \bigg(\frac{1}{\beta}\displaystyle\int_0^\beta O(x(\tau))\mathrm{d}\tau\bigg)\bigg]}{\mathbb{E}_\nu\bigg[\exp\bigg(-\displaystyle\int_0^\beta V^a(x(\tau))\mathrm{d}\tau\bigg)\bigg]}, \tag{3.100}$$

where ν is the Gaussian distribution defined in (3.93).

Proof The proof is accomplished in several steps. We establish the Trotter product formula

$$\lim_{N\to\infty} \langle q|\big(e^{-\frac{\beta_N}{2}V^a(\hat{q})}e^{-\beta_N\hat{H}^a}e^{-\frac{\beta_N}{2}V^a(\hat{q})}\big)^N|q\rangle = \langle q|e^{-\beta\hat{H}}|q\rangle\,, \tag{3.101}$$

and integrate over the variable $q \in \mathbb{R}^d$ to obtain

$$\mathrm{Tr}[e^{-\beta\hat{H}}O(\hat{q})] = \lim_{N\to\infty} \mathrm{Tr}\big[\big(e^{-\frac{\beta_N}{2}V^a(\hat{q})}e^{-\beta_N\hat{H}^a}e^{-\frac{\beta_N}{2}V^a(\hat{q})}\big)^N O(\hat{q})\big]. \tag{3.102}$$

Then we represent the RHS of (3.102) in the path integral form and calculate its limit as $N \to \infty$.

1. Simplification of the trace-expectation equality (3.100)

First, introduce the Hamiltonian operator of the quantum harmonic oscillator

$$\hat{H}^a = -\frac{\Delta_d}{2} + \frac{a^2}{2}|q|^2, \tag{3.103}$$

where Δ_d is the Laplace operator in $\mathbb{R}^d$. From Theorem X.29 of [39], we deduce that both $\hat{H}^a$ and $\hat{H}$ are essentially self-adjoint operators on the domain $C_0^\infty(\mathbb{R}^d)$, which comprises all smooth functions in $\mathbb{R}^d$ with compact support. Now we aim to prove

$$\frac{\mathrm{Tr}[e^{-\beta\hat{H}}O(\hat{q})]}{\mathrm{Tr}[e^{-\beta\hat{H}^a}]} = \mathbb{E}_\nu\bigg[\exp\bigg(-\int_0^\beta V^a(x(\tau))\mathrm{d}\tau\bigg)\times\bigg(\frac{1}{\beta}\int_0^\beta O(x(\tau))\mathrm{d}\tau\bigg)\bigg]. \tag{3.104}$$

In particular, by choosing $O(q)\equiv 1$ in the (3.104), we have

$$\frac{\mathrm{Tr}[e^{-\beta\hat{H}}]}{\mathrm{Tr}[e^{-\beta\hat{H}^a}]} = \mathbb{E}_\nu\bigg[\exp\bigg(-\int_0^\beta V^a(x(\tau))\mathrm{d}\tau\bigg)\bigg], \tag{3.105}$$

and from (3.104)(3.105) we immediately obtain (3.100). In the following we focus on proving (3.104).

2. Limit of $\langle q|\big(e^{-\frac{\beta_N}{2}V^a(\hat{q})}e^{-\beta_N\hat{H}^a}e^{-\frac{\beta_N}{2}V^a(\hat{q})}\big)^N|\tilde{q}\rangle$ as $N\to\infty$

Given the positions $q,\tilde{q}\in\mathbb{R}^d$, we aim to study the limit of the propagator

$$\langle q|\big(e^{-\frac{\beta_N}{2}V^a(\hat{q})}e^{-\beta_N\hat{H}^a}e^{-\frac{\beta_N}{2}V^a(\hat{q})}\big)^N|\tilde{q}\rangle \tag{3.106}$$

as $N\to\infty$. Observe that for any nonnegative function $\psi\in L^2(\mathbb{R}^d)$, $e^{-\tau\hat{H}^a}$ and $e^{-\tau V^a(\hat{q})}$ are both positivity-preserving operators, and using $V^a(q)\geqslant -M_1$ we have the inequalities

$$0\leqslant e^{-\tau\hat{H}^a}\psi(q)\leqslant\psi(q),\qquad 0\leqslant e^{-\tau V^a(\hat{q})}\psi(q)\leqslant e^{\tau M_1}\psi(q). \tag{3.107}$$

Hence we have the uniform-in-N upper bound

$$\begin{aligned}
&\langle q|\big(e^{-\frac{\beta_N}{2}V^a(\hat{q})}e^{-\beta_N\hat{H}^a}e^{-\frac{\beta_N}{2}V^a(\hat{q})}\big)^N|\tilde{q}\rangle \\
\leqslant\ &\langle q|\big(e^{\frac{\beta_N}{2}M_1}e^{-\beta_N\hat{H}^a}e^{\frac{\beta_N}{2}M_1}\big)^N|q\rangle = e^{\beta M_1}\langle q|e^{-\beta\hat{H}^a}|\tilde{q}\rangle \\
=\ &e^{\beta M_1}\bigg(\frac{a}{2\pi\sinh(a\beta)}\bigg)^{\frac{d}{2}}\exp\bigg(-\frac{a}{\sinh(a\beta)}\bigg(\cosh(a\beta)\frac{|q|^2+|\tilde{q}|^2}{2}-q\cdot\tilde{q}\bigg)\bigg),
\end{aligned} \tag{3.108}$$

where the explicit expression of $\langle q|e^{-\beta\hat{H}^a}|\tilde{q}\rangle$ is given by the Mehler kernel (see Theorem 1.5.10 of [24]). From (3.108) we conclude that (3.106) has a uniform-in-N upper bound.

Furthermore, we show that the propagator (3.106) can be represented in the form of the path integral. For given $q,\tilde{q}\in\mathbb{R}^d$, let $U_{q,\tilde{q}}^\beta$ be the Wiener measure of the continuous loop $x(\tau)\in\mathbb{H}$ defined by the following rule: for given $0<t_1<t_2<\cdots<t_{N-1}<\beta$, the measure of the set

$$\{x(\tau)\in\mathcal{H}: x(0)=q, x(\beta)=\tilde{q}, x(t_j)\in I_j, j=1,2,\cdots,N-1\} \tag{3.109}$$

is defined by

$$\int_{I_1}\mathrm{d}x_1\int_{I_2}\mathrm{d}x_2\cdots\int_{I_{N-1}}\mathrm{d}x_{N-1}\prod_{j=0}^{N-1}\langle x_j|e^{-(\tau_{j+1}-\tau_j)\hat{H}^a}|x_{j+1}\rangle\,, \tag{3.110}$$

where $I_1,I_2,\cdots,I_{N-1}$ are closed cuboids in $\mathbb{R}^d$, and we presume

$$t_0=0,\qquad x_0=q,\qquad t_N=\beta,\qquad x_N=\tilde{q}. \tag{3.111}$$

Similar to Equation (3.2.10) of [24], introduce the free positions $x_1, \cdots, x_{N-1}$ and write (3.106) as

$$
\begin{aligned}
&\langle q|\big(e^{-\frac{\beta_N}{2}V^a(\hat{q})}e^{-\beta_N\hat{H}^a}e^{-\frac{\beta_N}{2}V^a(\hat{q})}\big)^N|\tilde{q}\rangle\\
&=\int_{\mathbb{R}^d}\mathrm{d}x_1\cdots\int_{\mathbb{R}^d}\mathrm{d}x_{N-1}\prod_{j=0}^{N-1}\langle x_j|e^{-\frac{\beta_N}{2}V^a(\hat{q})}e^{-\beta_N\hat{H}^a}e^{-\frac{\beta_N}{2}V^a(\hat{q})}|x_{j+1}\rangle\\
&=\int_{\mathbb{R}^d}\mathrm{d}x_1\cdots\int_{\mathbb{R}^d}\mathrm{d}x_{N-1}\exp\Big(-\frac{\beta_N}{2}V^a(q)-\beta_N\sum_{j=1}^{N-1}V^a(x_j)-\frac{\beta_N}{2}V^a(\tilde{q})\Big)\prod_{j=0}^{N-1}\langle x_j|e^{-\beta_N\hat{H}^a}|x_{j+1}\rangle\\
&=\int\exp\Big(-\frac{\beta_N}{2}V^a(q)-\beta_N\sum_{j=1}^{N-1}V^a(x(j\beta_N))-\frac{\beta_N}{2}V^a(\tilde{q})\Big)\mathrm{d}U^\beta_{q,\tilde{q}},
\end{aligned}
\tag{3.112}
$$

where the integration is taken over the continuous loop $x(\tau)$ in the Wiener measure $U^\beta_{q,\tilde{q}}$. Let N tends to infinity in the equality (3.112), the dominated convergence theorem implies

$$
\lim_{N\to\infty}\langle q|\big(e^{-\frac{\beta_N}{2}V^a(\hat{q})}e^{-\beta_N\hat{H}^a}e^{-\frac{\beta_N}{2}V^a(\hat{q})}\big)^N|\tilde{q}\rangle=\int\exp\Big(-\int_0^\beta V^a(x(\tau))\mathrm{d}\tau\Big)\mathrm{d}U^\beta_{q,\tilde{q}}. \tag{3.113}
$$

Multiply (3.113) by a test function $\psi(\tilde{q})=\langle q|\psi\rangle\in L^2(\mathbb{R}^d)$ and integrate over the variable $\tilde{q}\in\mathbb{R}^d$,

$$
\begin{aligned}
&\lim_{N\to\infty}\int_{\mathbb{R}^d}\langle q|\big(e^{-\frac{\beta_N}{2}V^a(\hat{q})}e^{-\beta_N\hat{H}^a}e^{-\frac{\beta_N}{2}V^a(\hat{q})}\big)^N|\tilde{q}\rangle\,\langle\tilde{q}|\psi\rangle\,\mathrm{d}\tilde{q}\\
&\qquad=\Big\langle\int\exp\Big(-\int_0^\beta V^a(x(\tau))\mathrm{d}\tau\Big)\mathrm{d}U^\beta_{q,\tilde{q}},\psi(\tilde{q})\Big\rangle_{L^2(\mathbb{R}^d)}.
\end{aligned}
\tag{3.114}
$$

Here, $\langle\cdot,\cdot\rangle_{L^2(\mathbb{R}^d)}$ denotes the standard inner product $L^2(\mathbb{R}^d)$. Using the equality $1=\int_{\mathbb{R}^d}|\tilde{q}\rangle\,\langle\tilde{q}|$, (3.114) can be simplified as

$$
\begin{aligned}
&\lim_{N\to\infty}\langle q|\big(e^{-\frac{\beta_N}{2}V^a(\hat{q})}e^{-\beta_N\hat{H}^a}e^{-\frac{\beta_N}{2}V^a(\hat{q})}\big)^N|\psi\rangle\\
&\qquad=\Big\langle\int\exp\Big(-\int_0^\beta V^a(x(\tau))\mathrm{d}\tau\Big)\mathrm{d}U^\beta_{q,\tilde{q}},\psi(\tilde{q})\Big\rangle_{L^2(\mathbb{R}^d)}
\end{aligned}
\tag{3.115}
$$

Since $\hat{H}^a$ and $\hat{H}$ are both essentially self-adjoint operators in $L^2(\mathbb{R}^d)$, we can apply the Trotter product formula (Theorem VIII.31 of [39]) to derive the strong limit

$$
\lim_{N\to\infty}\big(e^{-\frac{\beta_N}{2}V^a(\hat{q})}e^{-\beta_N\hat{H}^a}e^{-\frac{\beta_N}{2}V^a(\hat{q})}\big)^N=e^{-\beta\hat{H}},\quad\text{in the strong sense.} \tag{3.116}
$$

In particular, for the test function $\psi\in L^2(\mathbb{R}^d)$, we have

$$
\lim_{N\to\infty}\langle q|\big(e^{-\frac{\beta_N}{2}V^a(\hat{q})}e^{-\beta_N\hat{H}^a}e^{-\frac{\beta_N}{2}V^a(\hat{q})}\big)^N|\psi\rangle=\langle q|e^{-\beta\hat{H}}|\psi\rangle\,,\quad\text{in the } L^2(\mathbb{R}^d)\text{ sense.} \tag{3.117}
$$

Combining the equalities (3.114) and (3.117), we obtain

$$
\langle q|e^{-\beta\hat{H}}|\psi\rangle=\Big\langle\int\exp\Big(-\int_0^\beta V^a(x(\tau))\mathrm{d}\tau\Big)\mathrm{d}U^\beta_{q,\tilde{q}},\psi(\tilde{q})\Big\rangle_{L^2(\mathbb{R}^d)} \tag{3.118}
$$

Since $\psi(\tilde{q})$ can be any test function in $L^2(\mathbb{R}^d)$, we deduce that for any $q, \tilde{q} \in \mathbb{R}^d$

$$\langle q|e^{-\beta \hat{H}}|\tilde{q}\rangle = \int \exp\bigg(- \int_0^\beta V^a(x(\tau))\mathrm{d}\tau \bigg) \mathrm{d}U_{q,\tilde{q}}^\beta. \tag{3.119}$$

Combining (3.113) and (3.119), we obtain the Trotter product formula

$$\lim_{N\to\infty} \langle q|\big(e^{-\frac{\beta_N}{2}V^a(\hat{q})}e^{-\beta_N \hat{H}^a}e^{-\frac{\beta_N}{2}V^a(\hat{q})}\big)^N|\tilde{q}\rangle = \langle q|e^{-\beta\hat{H}}|\tilde{q}\rangle\,. \tag{3.120}$$

In particular, by choosing $\tilde{q} = q$ we have

$$\lim_{N\to\infty} \langle q|\big(e^{-\frac{\beta_N}{2}V^a(\hat{q})}e^{-\beta_N \hat{H}^a}e^{-\frac{\beta_N}{2}V^a(\hat{q})}\big)^N|q\rangle = \langle q|e^{-\beta\hat{H}}|q\rangle\,. \tag{3.121}$$

Since the $\langle q|\big(e^{-\frac{\beta_N}{2}V^a(\hat{q})}e^{-\beta_N \hat{H}^a}e^{-\frac{\beta_N}{2}V^a(\hat{q})}\big)^N|q\rangle$ has a uniform-in-N upper bound given in (3.108), we can apply the dominated convergence theorem to deduce

$$\lim_{N\to\infty} \int_{\mathbb{R}^d} \langle q|\big(e^{-\frac{\beta_N}{2}V^a(\hat{q})}e^{-\beta_N \hat{H}^a}e^{-\frac{\beta_N}{2}V^a(\hat{q})}\big)^N|q\rangle\, O(q)\mathrm{d}q = \int_{\mathbb{R}^d} \langle q|e^{-\beta\hat{H}}|q\rangle\, O(q)\mathrm{d}q, \tag{3.122}$$

which is exactly equivalent to

$$\mathrm{Tr}[e^{-\beta\hat{H}}O(\hat{q})] = \lim_{N\to\infty} \mathrm{Tr}\big[\big(e^{-\frac{\beta}{2}V^a(\hat{q})}e^{-\beta \hat{H}^a}e^{-\frac{\beta}{2}V^a(\hat{q})}\big)^N O(\hat{q})\big]. \tag{3.123}$$

3. Expansion of $\mathrm{Tr}[e^{-\beta\hat{H}}O(\hat{q})]$ in the ring polymer distribution

Using the Trotter product formula (3.123), we can conveniently approximate the trace term $\mathrm{Tr}[e^{-\beta\hat{H}}O(\hat{q})]$ in the ring polymer distribution. In fact, by inserting the free positions $x_1, \cdots, x_N \in \mathbb{R}^d$, we have

$$\begin{aligned}
&\mathrm{Tr}[e^{-\beta\hat{H}}O(\hat{q})] \\
={}& \lim_{N\to\infty} \mathrm{Tr}\big[\big(e^{-\frac{\beta}{2}V^a(\hat{q})}e^{-\beta \hat{H}^a}e^{-\frac{\beta}{2}V^a(\hat{q})}\big)^N O(\hat{q})\big] \\
={}& \lim_{N\to\infty} \int_{\mathbb{R}^{dN}} \exp\bigg(- \beta_N \sum_{j=1}^N V^a(x_j)\bigg) \times O(x_1) \times \prod_{j=1}^N \langle x_j|e^{-\beta_N\hat{H}^a}|x_{j+1}\rangle \times \mathrm{d}x_1\cdots\mathrm{d}x_N.
\end{aligned} \tag{3.124}$$

Using the symmetry of the expression (3.124) in x_1, we obtain

$$\begin{aligned}
\mathrm{Tr}[e^{-\beta\hat{H}}O(\hat{q})] = \lim_{N\to\infty} \int_{\mathbb{R}^{dN}} & \exp\bigg(- \beta_N \sum_{j=1}^N V^a(x_j)\bigg) \times \\
& \bigg(\frac{1}{N}\sum_{j=1}^N O(x_j)\bigg) \times \prod_{j=1}^N \langle x_j|e^{-\beta_N\hat{H}^a}|x_{j+1}\rangle \times \mathrm{d}x_1\cdots\mathrm{d}x_N.
\end{aligned} \tag{3.125}$$

Therefore, we can define $\Theta_N^a(x_1, \cdots, x_N)$ to be the probability density of ring polymer part

$$\Theta_N^a(x_1, \cdots, x_N) = \frac{1}{Z^a} \prod_{j=1}^N \langle x_j|e^{-\beta\hat{H}^a}|x_{j+1}\rangle\,, \tag{3.126}$$

where each $\langle x_j | e^{-\beta \hat{H}^a} | x_{j+1} \rangle$ is given by the Mehler kernel, and Z^a is the normalization constant. By choosing $V^a(q) \equiv 0$ and $O(q) \equiv 1$ in (3.125) we have

$$\mathrm{Tr}[e^{-\beta \hat{H}^a}] = \lim_{N\to\infty} \int_{\mathbb{R}^{dN}} \prod_{j=1}^N \langle x_j | e^{-\beta \hat{H}^a} | x_{j+1} \rangle \times \mathrm{d}x_1 \cdots \mathrm{d}x_N. \tag{3.127}$$

Comparing (3.125) and (3.127) we immediately obtain

$$\begin{aligned} \frac{\mathrm{Tr}[e^{-\beta \hat{H}} O(\hat{q})]}{\mathrm{Tr}[e^{-\beta \hat{H}^a}]} = \lim_{N\to\infty} \int_{\mathbb{R}^{dN}} & \exp\bigg(-\beta_N \sum_{j=1}^N V^a(x_j) \bigg) \times \\ & \bigg(\frac{1}{N} \sum_{j=1}^N O(x_j) \bigg) \times \Theta_N^a(x_1, \cdots, x_N) \times \mathrm{d}x_1 \cdots \mathrm{d}x_N. \end{aligned} \tag{3.128}$$

Now define the Gaussian distribution $\tilde{U}^\beta$ in the Hilbert space $\mathbb{H} = L_p^2([0,\beta];\mathbb{R}^d)$ by the following rule: for any constants $0 \leqslant \tau_1 < \cdots < \tau_n < \beta$, the joint distribution of the random variables $x_1 = x(\tau_1), \cdots, x_n = x(\tau_n)$ is proportional to the density function

$$\prod_{j=1}^{n-1} \langle x_j | e^{-(\tau_{j+1}-\tau_j)\hat{H}^a} | x_{j+1} \rangle \times \langle x_n | e^{-(\tau_1 - \tau_n + \beta)\hat{H}^a} | x_1 \rangle, \tag{3.129}$$

which is product of Mehler kernels of the n adjacent pairs in the position coordinates $x(\tau_1), \cdots, x(\tau_n)$. From the Kolmogorov extension theorem, $\tilde{U}^\beta$ is indeed a well-defined distribution in $\mathbb{H}$. The difference between the Gaussian distribution $\tilde{U}^\beta$ and the Wiener measure $U_{q,q}^\beta$ defined in (3.110) is that the endpoints of the continuous loop in $\tilde{U}^\beta$ is flexible. Then $\Theta_N^a(x_1, \cdots, x_N)$ can be viewed as the joint distribution of the N positions coordinates $\{x(j\beta_N)\}_{j=1}^N$, where the continuous loop $x(\cdot)$ is randomly sampled from $\tilde{U}^\beta$. As a consequence, we can rewrite (3.128) as

$$\frac{\mathrm{Tr}[e^{-\beta \hat{H}} O(\hat{q})]}{\mathrm{Tr}[e^{-\beta \hat{H}^a}]} = \lim_{N\to\infty} \int \exp\bigg(-\beta_N \sum_{j=1}^N V^a(x(j\beta_N)) \bigg) \times \bigg(\frac{1}{N} \sum_{j=1}^N O(x(j\beta_N)) \bigg) \times \mathrm{d}\tilde{U}^\beta. \tag{3.130}$$

Using the dominated convergence theorem, (3.130) implies

$$\frac{\mathrm{Tr}[e^{-\beta \hat{H}} O(\hat{q})]}{\mathrm{Tr}[e^{-\beta \hat{H}^a}]} = \int \exp\bigg(-\int_0^\beta V^a(x(\tau))\mathrm{d}\tau \bigg) \times \bigg(\frac{1}{\beta} \int_0^\beta O(x(\tau))\mathrm{d}\tau \bigg) \times \mathrm{d}\tilde{U}^\beta. \tag{3.131}$$

4. Verification of the equivalence between the Gaussian distributions

Finally, we need to verify the Gaussian distribution $\tilde{U}^\beta$ is equivalent to the distribution ν defined in (3.93), which is expressed in the normal mode coordinates

$$x(\tau) = \sum_{k=0}^\infty \xi_k c_k(\tau), \quad \xi_k \sim \mathcal{N}\bigg(0, \frac{I_d}{\omega_k^2 + a^2}\bigg), \quad k = 0, 1, 2, \cdots. \tag{3.132}$$

In fact, $\tilde{U}^\beta$ and ν are both zero-mean Gaussian processes, and thus we only need to check their covariance functions are the same.

1. In the distribution $\tilde{U}^\beta$, the joint distribution of $x = x(0)$ and $y = x(\tau)$ is

$$\langle x|e^{-\tau \hat{H}^a}|y\rangle \, \langle y|e^{-(\beta-\tau)\hat{H}^a}|x\rangle = \left(\frac{a}{2\pi \sinh(a\beta)}\right)^d \exp\left(-\frac{A}{2}(|x|^2+|y|^2) + Bx^{\mathrm{T}}y\right), \quad (3.133)$$

   where the constants A, B are given by

$$A = \frac{a}{\tanh(a\tau)} + \frac{a}{\tanh(a(\beta-\tau))}, \qquad B = \frac{a}{\sinh(a\tau)} + \frac{a}{\sinh(a(\beta-\tau))}. \quad (3.134)$$

   Hence the covariance function is

$$\mathbb{E}_{\tilde{U}^\beta}[x(0)x(\tau)] = \frac{B}{A^2 - B^2}. \quad (3.135)$$

2. In the distribution ν, the covariance function is

$$\begin{aligned}\mathbb{E}_\nu[x(0)x(\tau)] &= \mathbb{E}_\nu\left[\sum_{k=0}^\infty \xi_k c_k(0) \sum_{k=0}^\infty \xi_k c_k(\tau)\right] = \sum_{k=0}^\infty \frac{c_k(0)c_k(\tau)}{\omega_k^2 + a^2} \\ &= \frac{1}{\beta}\cdot\frac{1}{a^2} + \frac{2}{\beta}\sum_{k=1}^\infty \frac{\cos\frac{2k\pi\tau}{\beta}}{(\frac{2k\pi}{\beta})^2 + a^2} = \frac{1}{\beta}\sum_{k\in\mathbb{Z}} \frac{\cos\frac{2k\pi\tau}{\beta}}{(\frac{2k\pi}{\beta})^2 + a^2}. \quad (3.136)\end{aligned}$$

Using Lemma A.1, the RHS of (3.135) and (3.136) are exactly the same. Therefore, we conclude that $\tilde{U}^\beta$ and ν are the same Gaussian distribution, which completes the proof. $\square$

Combining the results form Lemma 3.1, Lemma 3.2 and Theorem 3.3, we immediately obtain

Theorem 3.4 Under Assumptions (i)(iv), for any positive integer $N \in \mathbb{N}$,

$$|\langle O(\hat{q})\rangle_{\beta,N} - \langle O(\hat{q})\rangle_\beta| \leqslant \frac{d\beta M_4}{8N} + 2\sqrt{d(\beta+2C_0)}\beta M_1 M_4 \exp\left(\frac{5}{2}\beta M_1 + C_0 M_1\right)\frac{1}{\sqrt{N}}. \quad (3.137)$$

Remark If the potential function $V^a(q) = 0$, the constants M_1 can be chosen as 0 and the convergence rate become $\mathcal{O}(1/N)$ instead of $\mathcal{O}(1/\sqrt{N})$.

Using the geometric ergodicity results previously, we can further obtain the sampling error of the overdamped Langevin dynamics (3.6) and the underdamped Langevin dynamics (3.7). For the overdamped Langevin dynamics (3.6), define the statistical average at time t by

$$\langle O(\hat{q})\rangle_{\beta,N,t} = \int_{\mathbb{R}^{dN}} \left[\frac{1}{\beta}\int_0^\beta O\bigg(\sum_{k=0}^{N-1}\xi_k c_k(\tau)\bigg)\mathrm{d}\tau\right](\nu_N P_t)(\xi)\mathrm{d}\xi, \quad (3.138)$$

then using the Pinsker's inequality,

$$\begin{aligned}|\langle O(\hat{q})\rangle_{\beta,N,t} - \langle O(\hat{q})\rangle_{\beta,N}| &\leqslant M_4 \int_{\mathbb{R}^{dN}} |(\nu_N P_t)(\xi) - \pi_N(\xi)|\mathrm{d}\xi \\ &\leqslant M_4\sqrt{2H(\nu_N P_t|\pi_N)} \leqslant \sqrt{2C_1}M_4\exp(-\lambda_1 t).\end{aligned}$$

Hence from (3.91) we have

$$|\langle O(\hat{q})\rangle_{\beta,N,t} - \langle O(\hat{q})\rangle_\beta| \leqslant \frac{K_1}{N} + \frac{K_2}{\sqrt{N}} + \sqrt{2C_1}M_4 \exp(-\lambda_1 t), \tag{3.139}$$

where the constants K_1 and K_2 are defined by

$$K_1 = \frac{d\beta M_4}{8}, \qquad K_2 = 2\sqrt{d(\beta+2C_0)}\beta M_1 M_4 \exp\left(\frac{5}{2}\beta M_1 + C_0 M_1\right). \tag{3.140}$$

Similarly, for the underdamped Langevin dynamics (3.7) we define

$$\langle O(\hat{q})\rangle_{\beta,N,t} = \int_{\mathbb{R}^{dN}\times\mathbb{R}^{dN}} \left[\frac{1}{\beta}\int_0^\beta O\bigg(\sum_{k=0}^{N-1}\xi_k c_k(\tau)\bigg)\mathrm{d}\tau\right](\sigma_N P_t)(\xi,\eta)\mathrm{d}\xi, \tag{3.141}$$

then

$$|\langle O(\hat{q})\rangle_{\beta,N,t} - \langle O(\hat{q})\rangle_\beta| \leqslant \frac{K_1}{N} + \frac{K_2}{\sqrt{N}} + \sqrt{2C_2}M_4 \exp(-\lambda_2 t). \tag{3.142}$$

The results above can be summarized as

Corollary 3.3 For the overdamped Langevin dynamics (3.6), under Assumptions (i)(ii)(iv), the difference between the statistical average $\langle O(\hat{q})\rangle_{\beta,N,t}$ and the quantum thermal average $\langle O(\hat{q})\rangle_\beta$ is estimated by

$$\big|\langle O(\hat{q})\rangle_{\beta,N,t} - \langle O(\hat{q})\rangle_\beta\big| \leqslant \frac{K_1}{N} + \frac{K_2}{\sqrt{N}} + \sqrt{2C_1}M_4 \exp(-\lambda_1 t), \qquad \forall t \geqslant 0. \tag{3.143}$$

For the underdamped Langevin dynamics (3.7), under Assumptions (i)(ii)(iii)(iv), the difference between the statistical average $\langle O(\hat{q})\rangle_{\beta,N,t}$ and the quantum thermal average $\langle O(\hat{q})\rangle_\beta$ is estimated by

$$\big|\langle O(\hat{q})\rangle_{\beta,N,t} - \langle O(\hat{q})\rangle_\beta\big| \leqslant \frac{K_1}{N} + \frac{K_2}{\sqrt{N}} + \sqrt{2C_2}M_4 \exp(-\lambda_2 t), \qquad \forall t \geqslant 0. \tag{3.144}$$

The constants K_1, K_2 are given in (3.140), and the constants C_1, C_2 are given in Corollary 3.1 and Corollary 3.2, respectively.

4 Numerical method for the CL-PIMD

We elaborate on the numerical method for the truncated CL-PIMD (3.1). For given $N \in \mathbb{N}$, the numerical method for the truncated CL-PIMD (3.1) consists of two parts:

1. For given normal mode coordinates $\{\xi_k\}_{k=0}^{N-1}$, compute the integral terms

$$\int_0^\beta \nabla V^a(x_N(\tau))c_k(\tau)\mathrm{d}\tau \quad \text{and} \quad \int_0^\beta O(x_N(\tau))\mathrm{d}\tau, \tag{4.1}$$

numerically for $k = 0, 1, \cdots, N-1$, where $x_N(\tau) = \sum_{k=0}^{N-1}\xi_k c_k(\tau)$ is the continuous loop.

2. Integrate the underdamped Langevin dynamics (3.7) numerically.

For general forms of the functions $V^a(x)$ and $O(x)$, the integral terms in (4.1) cannot be computed analytically, hence we employ the numerical integration to approximate these integral terms. Picking an integer $D \in \mathbb{N}$, let the step size be $\beta_D = \beta/D$ and compute (4.1) from

$$\beta_D \sum_{j=0}^{D-1} \nabla V^a(x_N(j\beta_D))c_k(j\beta_D) \quad \text{and} \quad \beta_D \sum_{j=0}^{D-1} O(x(j\beta_D)). \tag{4.2}$$

Since $\{c_k(\tau)\}_{k=0}^{\infty}$ are all trigonometric functions, the summation in (4.2) can be conveniently computed by the fast Fourier transform (FFT).

Next we demonstrate the BAOAB scheme for the simulation of the truncated CL-PIMD (3.1). The BAOAB scheme is a widely used numerical integrator in the molecular dynamics [40] and can be made strongly stable in the case of the standard PIMD [23,33]. We show that the strongly stable BAOAB scheme can also be applied in the truncated CL-PIMD (3.1). To begin with, decompose the Langevin dynamics in the truncated CL-PIMD (3.1) into three parts:

$$A : \begin{cases} \dot{\xi}_k = \eta_k, \\ \dot{\eta}_k = -\xi_k, \end{cases} \qquad \text{(evolution of the potential-free continuous loop)}$$

$$B : \dot{\eta}_k = -\frac{1}{\omega_k^2 + a^2} \int_0^\beta \nabla V^a(x_N(\tau))c_k(\tau)\mathrm{d}\tau, \qquad \text{(drift force from potential function)}$$

$$C : \dot{\eta}_k = -\gamma_k \eta_k + \sqrt{\frac{2\gamma_k}{\omega_k^2 + a^2}} \dot{B}_k. \qquad \text{(Ornstein–Uhlenbeck process)}$$

The part A for each mode can be viewed as a linear system with the coefficient matrix

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \tag{4.3}$$

and the corresponding strongly stable numerical integrator is given by the Cayley modification

$$\mathrm{cay}(\Delta t A) = \Big(I_2 - \frac{\Delta t}{2} A\Big)^{-1} \Big(I_2 + \frac{\Delta t}{2} A\Big). \tag{4.4}$$

The part B and part O are explicitly solved as

$$\exp(\Delta t B) : \eta_k(\Delta t) = \eta_k(0) - \frac{\Delta t}{\omega_k^2 + a^2} \int_0^\beta \nabla V^a(x_N(\tau))c_k(\tau)\mathrm{d}\tau, \tag{4.5}$$

$$\exp(\Delta t O) : \eta_k(\Delta t) = e^{-\gamma_k \Delta t}\eta_k(0) + \sqrt{\frac{1 - e^{-2\gamma_k \Delta t}}{\omega_k^2 + a^2}} \theta_k, \tag{4.6}$$

where $\theta_k \sim \mathcal{N}(0, I_d)$ is a Gaussian random variable independently sampled at each time step. Finally, the BAOAB scheme with Cayley modification is given by

$$\mathrm{BAOAB} : \exp\Big(\frac{\Delta t}{2} B\Big) \mathrm{cay}\Big(\frac{\Delta t}{2} A\Big) \exp(\Delta t O) \mathrm{cay}\Big(\frac{\Delta t}{2} A\Big) \exp\Big(\frac{\Delta t}{2} B\Big). \tag{4.7}$$

In the numerical experiments, we simply choose $\gamma_k = 1$ for all normal modes so that the time step Δt can be chosen to be moderately small to ensure the stability of the numerical simulation.

In conclusion, the computational cost of the truncated CL-PIMD (3.1) per time step is

$$\text{cost per time step} = \min\{N, D\} \log D + (c_{\nabla V} + c_O) D, \tag{4.8}$$

where $c_{\nabla V}$ and c_O are the computational cost of the gradient function $\nabla V(q)$ and the observable function $O(q)$, respectively. When the number of grid points $D = N$ and N is an odd integer, the resulting truncated CL-PIMD (3.1) becomes exactly the standard PIMD with preconditioning, which is also referred to as the pmmLang in [22]. Therefore, the standard PIMD can be viewed as a special case of the truncated CL-PIMD (3.1) with the choice $D = N$.

In the numerical simulation of the truncated CL-PIMD (3.1), one need to specify the two parameters N and D carefully and find a balance between the computational cost and the sampling accuracy. Based on the computational cost derived in (4.8), if the calculation of the functions $\nabla V(q)$ or $O(q)$ is expensive, we can choose a relatively large N with respect to D. Such modification of the standard PIMD ($D = N$) is expected to achieve better accuracy without increasing too much complexity. See also the numerical experiments in Section 5. The idea of increasing the number of modes N with respect to given D is also similar to the ring polymer contraction described in [20].

5 Numerical results of the CL-PIMD

Consider the 1D quantum system with the potential function

$$V(q) = \frac{1}{2}q^2 + q\cos q, \qquad q \in \mathbb{R}, \tag{5.1}$$

and we aim to compute $\langle O(\hat{q})\rangle_\beta$ using the CL-PIMD, where the observable function

$$O(q) = \sin\frac{\pi q}{2} \qquad q \in \mathbb{R}. \tag{5.2}$$

The graph of the potential function $V(q)$ and the observable functions are shown in Figure 2.

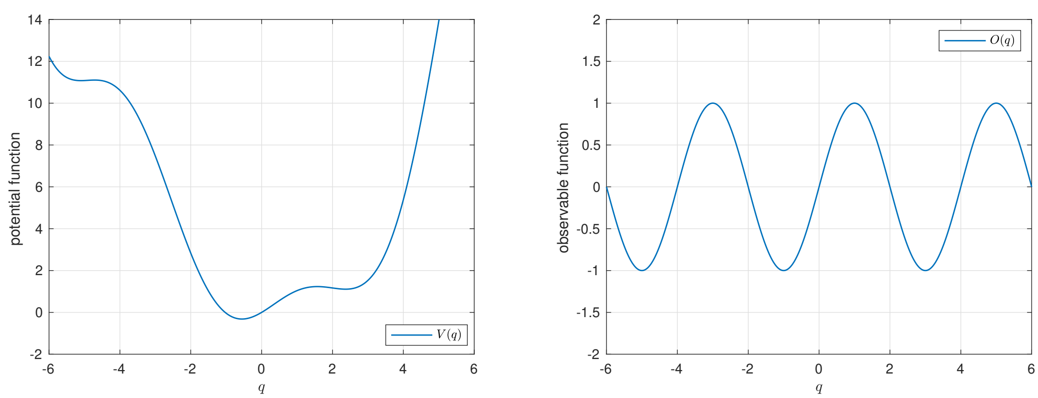

Figure 2: Left: potential function $V(q)$. Right: observable function $O(q)$.

Choose the constant $a = 1$. Let the number of normal modes be a positive integer $N \in \mathbb{N}$, and define the loop integral of the observable function $O(q)$ by

$$\bar{O}_N(\xi) = \frac{1}{\beta}\int_0^\beta O\bigg(\sum_{k=0}^{N-1}\xi_k c_k(\tau)\bigg)\mathrm{d}\tau. \tag{5.3}$$

Suppose the normal mode coordinates $\{\xi^t\}_{t\geqslant 0}$ and $\{\eta^t\}_{t\geqslant 0}$ are evolved by the truncated CL-PIMD (3.1), and the initial values of ξ^0 and η^0 are chosen from the Gaussian distribution σ_N defined by

$$\xi_k^0, \eta_k^0 \sim \mathcal{N}\bigg(0, \frac{I_d}{\omega_k^2 + a^2}\bigg), \qquad k = 0, 1, \cdots, N-1. \tag{5.4}$$

Choose the damping rates $\gamma_k = 1$ for all normal modes, then we use the following time average

$$\frac{1}{T}\int_0^T \bar{O}_N(\xi^t)\mathrm{d}t \tag{5.5}$$

to approximate the statistical average $\langle O(q)\rangle_{\beta,N}$ defined in (3.5) as $T \to \infty$. Finally, as the number of modes $N \to \infty$, the time average (5.5) is expected to be a reasonable approximation of $\langle O(\hat{q})\rangle_\beta$. By propagating a single trajectory of $\{\xi^t\}_{t\geqslant 0}$, the discrepancy of the time average (5.5) from $\langle O(\hat{q})\rangle_\beta$ can thus be used to characterize the convergence properties of the truncated CL-PIMD (2.25).

Define the observable error $e_{N,T}$ by

$$e_{N,T} = \bigg|\frac{1}{T}\int_0^T \bar{O}_N(\xi^t)\mathrm{d}t - \langle O(\hat{q})\rangle_\beta\bigg|, \tag{5.6}$$

then the decay of $e_{N,T}$ with respect to the number of normal modes N reflects the convergence rate of the truncated CL-PIMD (3.1). Next, define the time correlation function $c_{N,T}$ by

$$c_{N,T} = \mathbb{E}\bigg[\Big(\bar{O}_N(\xi^0) - \langle O(\hat{q})\rangle_{\beta,N}\Big)\Big(\bar{O}_N(\xi^T) - \langle O(\hat{q})\rangle_{\beta,N}\Big) : \xi^0 \sim \pi_N(\xi)\bigg], \tag{5.7}$$

then the decay of $c_{N,T}$ can be used to characterize the long-time convergence rate of the truncated CL-PIMD (3.1). Note that in (5.7) we apply $\langle O(\hat{q})\rangle_{\beta,N}$ rather than $\langle O(\hat{q})\rangle_\beta$ because $\langle O(\hat{q})\rangle_{\beta,N}$ is the statistical average of $\bar{O}_N(\xi)$ in the distribution $\pi_N(\xi)$.

Remark In general, the convergence rate in the relative entropy is difficult to be computed from the data of a single trajectory, while the decay rate of the correlation function can be computed conveniently. Nevertheless, both the entropic convergence rate and the decay rate reflect the long-time convergence of a Markov process. Also, the decay of the correlation function can derived from the decay rate of the variance function (see Section 4.2 of [26]). In the case of the overdamped Langevin dynamics, the log-Sobolev inequality implies the Poincaré inequality, and thus the decay rate of correlation function is no smaller than the entropic convergence rate. In more complicated systems such as the degenerate diffusion processes, the relations between these rates need in-depth argumentation, which is not the main focus on this paper.

In the numerical setting, we evolve the truncated CL-PIMD (3.1) and compute the correlation function using the numerical integration and the BAOAB scheme, where the number of grid points

D for numerical integration is chosen as $D = 128$, the time step Δt in the BAOAB scheme is $\Delta t = 1/16$, and the total simulation time is $T = 10^6$. The reference values of the quantum thermal average is computed from the spectral method (see Appendix D). When the inverse temperature $\beta = 1, 2$, the reference value of $\langle O(\hat{q})\rangle_\beta$ is -0.2364 and -0.2932, respectively.

Remark The spectral method encounters the curse of dimensionality when the spatial dimension d is large, that is, the computational cost of the spectral method grows exponentially with d. On the contrary, the computational cost of the standard PIMD and the CL-PIMD in a time step grows linearly with d.

We first visualize the invariant distribution $\pi_N(\xi)$ by estimating its marginal distributions in the normal mode coordinates $\{\xi_k\}_{k=0}^{N-1}$. When the number of modes $N = 17$, we plot the marginal distributions of the initial distribution σ_N in (5.4) and the invariant distribution $\pi_N(\xi)$ in (3.4) in Figure 3 and Figure 4. Here, Figure 3 and Figure 4 correspond to the case $\beta = 1$ and $\beta = 2$, respectively. The left panel shows the distribution of $\{\xi_k\}_{k=0}^4$ in $\sigma_N(\xi)$, while the right panel shows the distribution of $\{\xi_k\}_{k=0}^4$ in $\pi_N(\xi)$. The red line and the black rectangle show the mean value and the standard deviation in each marginal distribution. Note that the first normal mode coordinate ξ_0 corresponds to the center of the continuous loop, because

$$\xi_0 = \int_0^\beta x(\tau)c_0(\tau)\mathrm{d}\tau = \sqrt{\frac{1}{\beta}}\int_0^\beta x(\tau)\mathrm{d}\tau. \tag{5.8}$$

Also, note that the integral of $V^a(x(\tau))$ has translation invariance, i.e.,

$$\int_0^\beta V^a\bigg(\sum_{k=0}^\infty \xi_k c_k(\tau)\bigg)\mathrm{d}\tau = \int_0^\beta V^a\bigg(\sum_{k=0}^\infty \xi_k c_k(\tau+\theta)\bigg)\mathrm{d}\tau, \qquad \forall\theta\in\mathbb{R}, \tag{5.9}$$

we deduce that the marginal distributions of the normal mode coordinates ξ_{2k-1} and ξ_{2k} are the same and symmetric with respect to the vertical axis. In particular, ξ_1 and ξ_2 have the same marginal distribution, while ξ_3 and ξ_4 have the same marginal distribution.

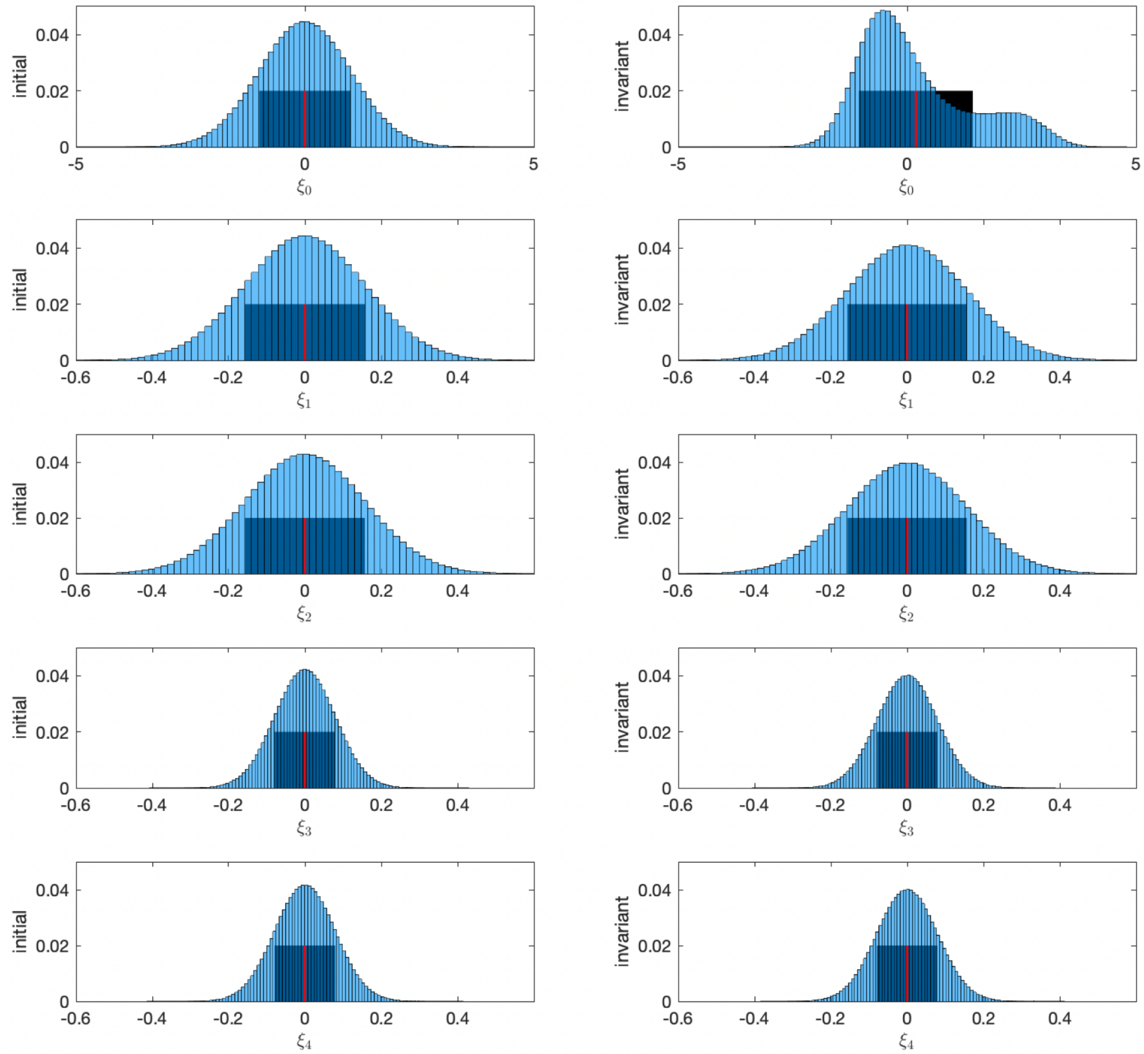

Figure 3: Marginal distributions of the normal mode coordinates $\{\xi_k\}_{k=0}^{4}$ in the initial distribution $\sigma_N(\xi)$ (left) and the invariant distribution $\pi_N(\xi)$ (right). The inverse temperature $\beta = 1$, and the number of normal modes $N = 17$.

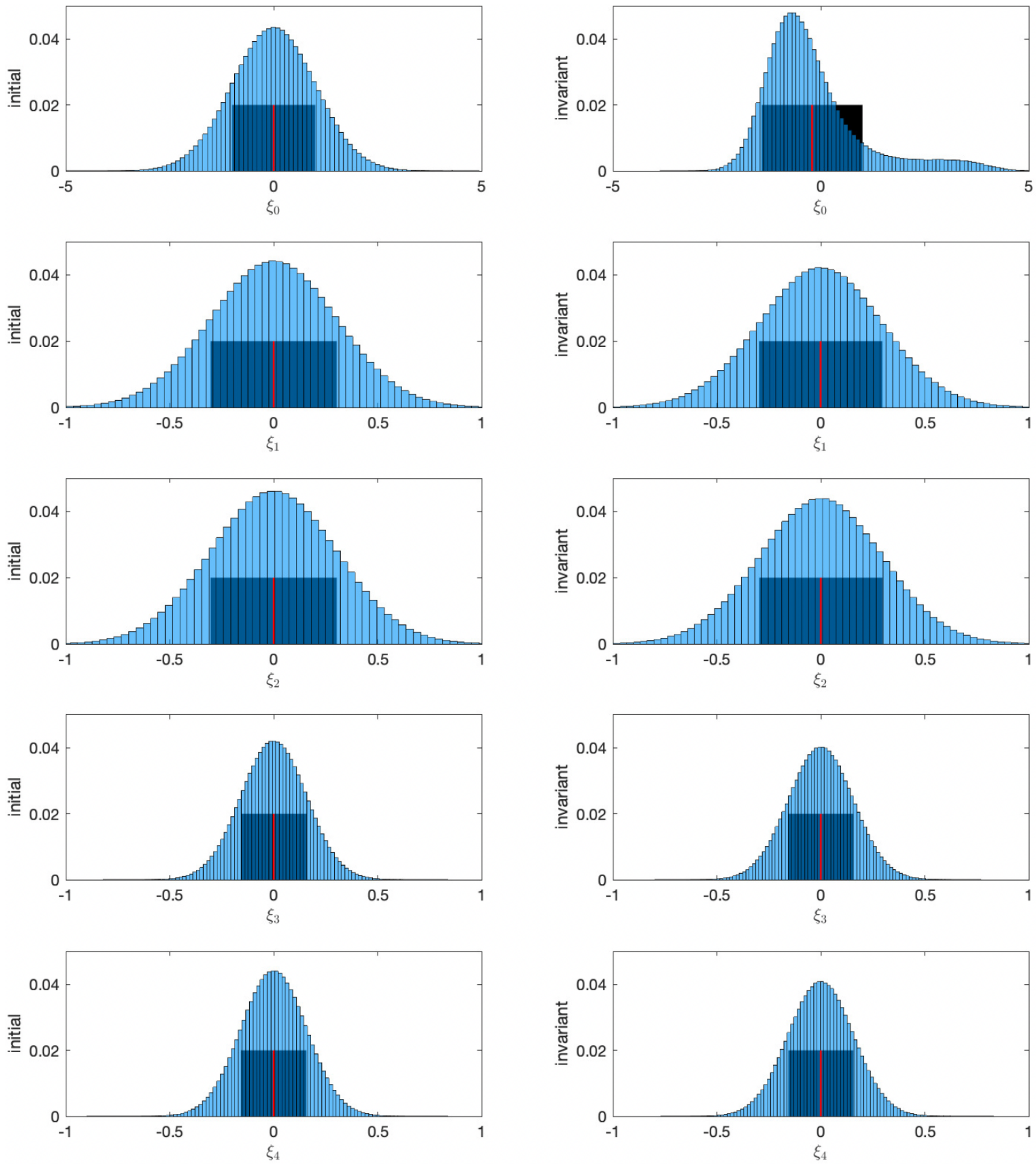

Figure 4: Marginal distributions of the normal mode coordinates $\{\xi_k\}_{k=0}^4$ in the initial distribution $\sigma_N(\xi)$ (left) and the invariant distribution $\pi_N(\xi)$ (right). The inverse temperature $\beta = 2$, and the number of normal modes $N = 17$.

Figure 3 and Figure 4 partially show how the invariant distribution $\pi_N(\xi)$ differs from the initial distribution $\sigma_N(\xi)$. Under the influence of the potential function $V^a(q)$, the marginal distribution of ξ_0 is similar to the classical Boltzmann distribution $e^{-\beta V(q)}$. In comparison, the change in the marginal distributions of $\{\xi_k\}_{k=1}^4$ are minor.

When the number of normal modes N varies in $\{5, 9, 17, 33, 65\}$, we plot the graphs of $e_{N,T}$ and $c_{N,T}$ along the simulation in Figure 5, Figure 6 and Figure 7. Note that in Figure 7 the absolute value of $c_{N,T}$ is shown in the logarithm form, and we draw the approximate linear bounds for $\log_2 |c_{N,T}|$. When the number of grid points D varies in $\{5, 9, 17, 33, 65\}$, we show the values of $e_{N,T}$ at the end of the simulation in Table 4.

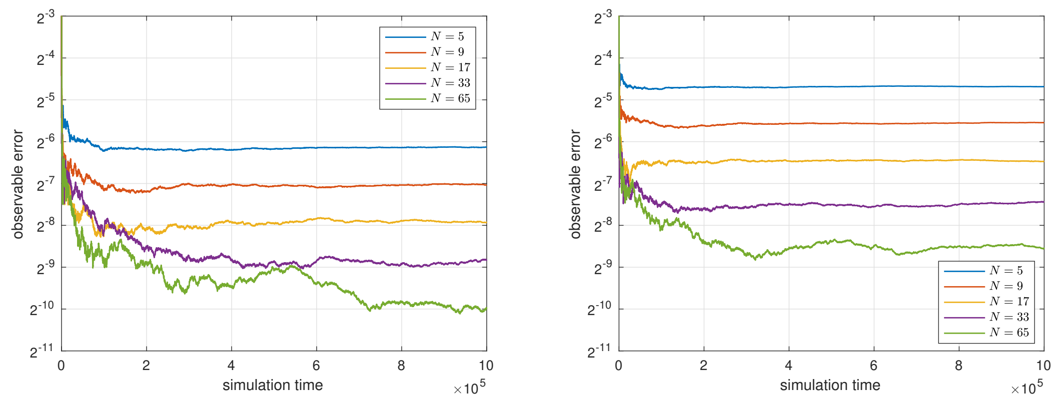

Figure 5: Observable error $e_{N,T}$ foxr different numbers of normal modes N in the simulation of the CL-PIMD. The left and right figures correspond to $\beta = 1$ and $\beta = 2$, respectively.

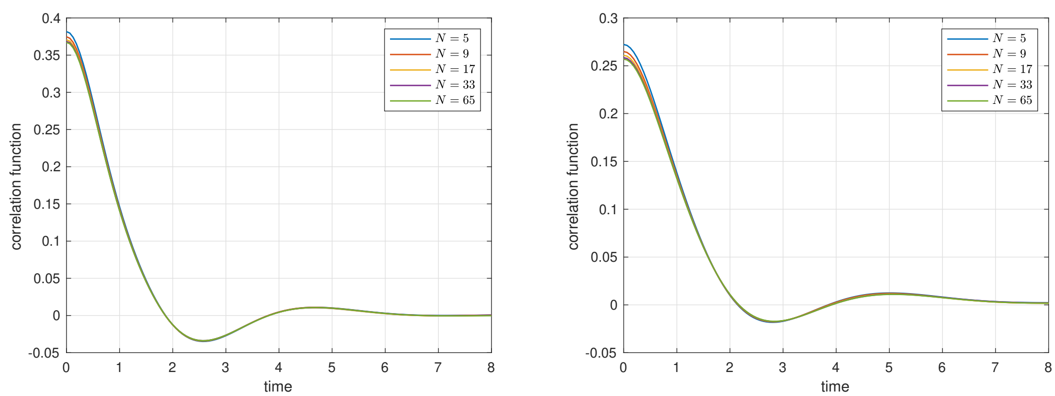

Figure 6: Correlation function $c_{N,T}$ for different numbers of normal modes N in the simulation of the CL-PIMD. The left and right figures correspond to $\beta = 1$ and $\beta = 2$, respectively. The x-axis denotes the time interval to calculate the correlation function rather than the real simulation time.

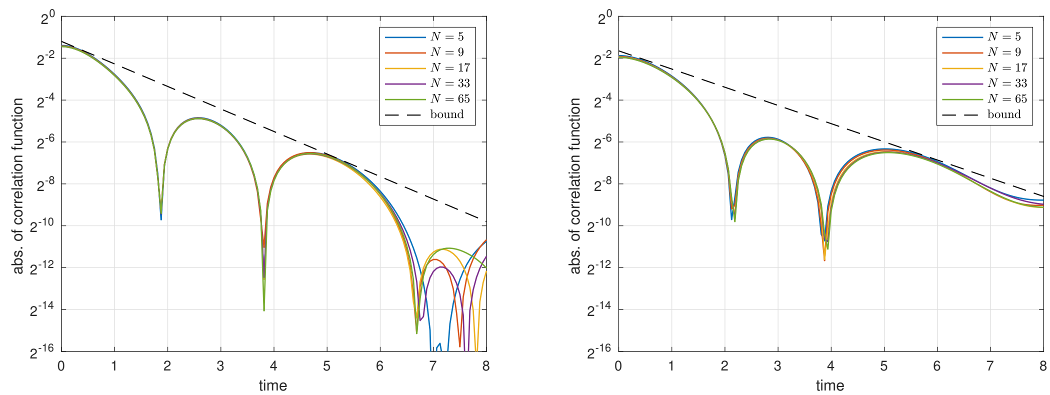

Figure 7: Absolute value of the correlation function $c_{N,T}$ for different numbers of normal modes N in the simulation of the CL-PIMD, shown in the logarithm form. The left and right figures correspond to $\beta = 1$ and $\beta = 2$, respectively.

$e_{N,T}$ $(\beta = 1)$	$N = 5$	$N = 9$	$N = 17$	$N = 33$	$N = 65$
$D = 5$	$\mathbf{1.431 \times 10^{-2}}$	7.896×10^{-3}	6.387×10^{-3}	5.206×10^{-3}	4.165×10^{-3}
$D = 9$	1.428×10^{-2}	$\mathbf{7.626 \times 10^{-3}}$	4.272×10^{-3}	2.961×10^{-3}	1.761×10^{-3}
$D = 17$	1.428×10^{-2}	7.624×10^{-3}	$\mathbf{4.119 \times 10^{-3}}$	2.300×10^{-3}	1.185×10^{-3}
$D = 33$	1.428×10^{-2}	7.624×10^{-3}	4.115×10^{-3}	$\mathbf{2.189 \times 10^{-3}}$	9.709×10^{-4}
$D = 65$	1.428×10^{-2}	7.624×10^{-3}	4.115×10^{-3}	2.187×10^{-3}	$\mathbf{9.867 \times 10^{-4}}$
$e_{N,T}$ $(\beta = 2)$	$N = 5$	$N = 9$	$N = 17$	$N = 33$	$N = 65$
$D = 5$	$\mathbf{3.900 \times 10^{-2}}$	2.344×10^{-2}	1.835×10^{-2}	1.517×10^{-2}	1.138×10^{-2}
$D = 9$	3.890×10^{-2}	$\mathbf{2.141 \times 10^{-2}}$	1.189×10^{-2}	8.026×10^{-3}	5.481×10^{-3}
$D = 17$	3.890×10^{-2}	2.141×10^{-2}	$\mathbf{1.127 \times 10^{-2}}$	5.892×10^{-3}	3.331×10^{-3}
$D = 33$	3.890×10^{-2}	2.141×10^{-2}	1.126×10^{-2}	$\mathbf{5.754 \times 10^{-3}}$	2.671×10^{-3}
$D = 65$	3.890×10^{-2}	2.141×10^{-2}	1.126×10^{-2}	5.750×10^{-3}	$\mathbf{2.678 \times 10^{-3}}$

Table 4: Observable error $e_{N,T}$ for different numbers of normal modes N and number of grid points D in the simulation of the CL-PIMD. The upper and lower parts correspond to the inverse temperature $\beta = 1$ and $\beta = 2$, respectively. Data on diagonal lines (marked in bold) are equivalent to the results of the standard PIMD.

It can be inferred from Figure 5, Figure 6 and Figure 7 that

1. The convergence order of the statistical average $\langle O(\hat{q}) \rangle_{\beta,N}$ towards the quantum thermal average $\langle O(\hat{q}) \rangle_{\beta}$ with respect to the number of normal modes N is approximately 1. When we double the parameter N, the observable error $e_{N,T}$ reduces to approximately half. In comparison, we prove that the convergence order is at least $1/2$ theoretically.
2. The correlation function $c_{N,T}$ decays exponentially to zero and decay rate does not depend on the number of normal modes N.
3. The observable error $e_{N,T}$ of the truncated CL-PIMD grows large and the convergence rate towards the equilibrium becomes slow when the inverse temperature β is large.

These numerical results substantially conform to the theoretical results in Section 3.

From Table 4 we observe that the observable error $e_{N,T}$ depends on both the number of normal modes N and the number of grid points D. Note that for fixed $D \in \mathbb{N}$, increasing N continues reduces the value of $e_{N,T}$ even when $N > D$. Therefore, if the calculation of $\nabla V(q)$ or $O(q)$ is expensive, we can appropriately choose a large integer N which is larger than D to ensure better accuracy. See also our discussion of the computational cost in Section 4.

6 Conclusion

We derive the continuous loop path integral molecular dynamics (CL-PIMD) from the infinite bead limit of the standard path integral representation. By truncating the number of normal modes to a finite integer $N \in \mathbb{N}$, we obtain the truncated CL-PIMD. We quantify the approximation error of the truncated CL-PIMD from the true quantum thermal average, and prove that the truncated CL-PIMD has uniform-in-N geometric ergodicity. These results are also supported by our numerical experiments. From the numerical perspective, choosing a large N relative to D enhances the accuracy of the truncated CL-PIMD. There are still a few open questions to study: How to define the CL-PIMD in the multi-level quantum system? Is the CL-PIMD compatible with other thermostats, like the Nosé–Hoover thermostat? How to apply the CL-PIMD to compute the quantum correlation function, which involves the real-time quantum dynamics?

A Additional proofs for Sections 3.2 & 3.3

Lemma A.1 Given the constants $a > 0$ and $\tau \in [0, 2\pi]$, the following equality holds true

$$\sum_{k\in\mathbb{Z}} \frac{\cos k\tau}{k^2 + a^2} = \frac{\pi \cosh a(\pi - \tau)}{a \sinh \pi a}. \tag{A.1}$$

Proof The proof stems from the user FShrike's answer on the problem *Calculate* $\sum_{n=1}^{\infty} \frac{a\cos(nx)}{a^2+n^2}$ in the website StackExchange: Mathematics. Define the complex-valued function

$$f(z) = \frac{e^{i\tau z}}{e^{2i\tau z} - 1} \cdot \frac{1}{z^2 + a^2}, \qquad z \in \mathbb{C}. \tag{A.2}$$

Then it is easy to observe that $\lim_{|z|\to\infty} z f(z) = 0$, and $f(z)$ has simple poles at $\pm ia$ and $\mathbb{Z}$. Let $N \in \mathbb{N}$ be a large integer, and integrate $f(z)$ along the box contors with vertices $\pm(N + \frac{1}{2})(1 \pm i)$, then the residue theorem implies

$$2\pi i \sum_{k\in\mathbb{Z}} \mathrm{Res}(f; k) = -2\pi i \big(\mathrm{Res}(f; ia) + \mathrm{Res}(f; -ia)\big). \tag{A.3}$$

The LHS of (A.3) is exactly $\sum_{k\in\mathbb{Z}} \frac{e^{ik\tau}}{k^2+a^2}$, and the RHS of (A.3) equals

$$-2\pi i \left(\frac{e^{-\tau a}}{2ia(e^{-2\pi a} - 1)} - \frac{e^{\tau a}}{2ia(e^{2\pi a - 1})} \right) = \frac{\pi}{a} \cdot \frac{\sinh(\tau a) + \sinh((2\pi - \tau)a)}{\cosh(2\pi a) - 1}, \tag{A.4}$$

and thus we obtain the equality (A.1). $\square$

Lemma A.2 Under Assumption (i), for any real sequences $\xi = \{\xi_k\}_{k=0}^{\infty}$ and $\eta_k = \{\eta_k\}_{k=0}^{\infty}$ in $\mathbb{R}^d$,

$$\begin{aligned}\bigg|\int_0^\beta V^a\Big(\sum_{k=0}^{\infty}\xi_k c_k(\tau)\Big)\mathrm{d}\tau - \int_0^\beta V^a\Big(\sum_{k=0}^{\infty}\eta_k c_k(\tau)\Big)\mathrm{d}\tau\bigg| \\ \leqslant \sqrt{3}M_1\sqrt{\sum_{k=0}^{\infty}|\xi_k-\eta_k|^2\Big(\beta+\sum_{k=0}^{\infty}|\xi_k|^2+\sum_{k=0}^{\infty}|\eta_k|^2\Big)}.\end{aligned} \tag{A.5}$$

Proof Note that for any $q_1, q_2 \in \mathbb{R}^d$, by Assumption (i) we have

$$\begin{aligned}|V^a(q_1)-V^a(q_2)| &\leqslant |q_1-q_2|\int_0^1 \big|\nabla V^a\big(q_1+\theta(q_2-q_1)\big)\big|\mathrm{d}\theta \\ &\leqslant |q_1-q_2|\Big(M_1+M_1\max_{\theta\in[0,1]}\big|q_1+\theta(q_2-q_1)\big|\Big) \\ &\leqslant M_1|q_1-q_2|\big(1+|q_1|+|q_2|\big) \\ &\leqslant \sqrt{3}M_1|q_1-q_2|\sqrt{1+|q_1|^2+|q_2|^2}.\end{aligned} \tag{A.6}$$

Now choose

$$q_1 = \sum_{k=0}^{\infty}\xi_k c_k(\tau), \qquad q_2 = \sum_{k=0}^{\infty}\eta_k c_k(\tau), \qquad \tau\in[0,\beta], \tag{A.7}$$

then

$$\begin{aligned}\bigg|V^a\Big(\sum_{k=0}^{\infty}\xi_k c_k(\tau)\Big) - V^a\Big(\sum_{k=0}^{\infty}\eta_k c_k(\tau)\Big)\bigg| &\leqslant \sqrt{3}M_1\bigg|\sum_{k=0}^{\infty}(\xi_k-\eta_k)c_k(\tau)\bigg| \\ &\quad\times\sqrt{1+\Big(\sum_{k=0}^{\infty}\xi_k c_k(\tau)\Big)^2+\Big(\sum_{k=0}^{\infty}\eta_k c_k(\tau)\Big)^2}.\end{aligned}$$

Integrating both sides in the interval $\tau\in[0,\beta]$, we obtain from the Cauchy's inequality,

$$\begin{aligned}&\int_0^\beta \bigg|V^a\Big(\sum_{k=0}^{\infty}\xi_k c_k(\tau)\Big) - V^a\Big(\sum_{k=0}^{\infty}\eta_k c_k(\tau)\Big)\bigg|\mathrm{d}\tau \\ &\leqslant \sqrt{3}M_1\sqrt{\int_0^\beta\bigg|\sum_{k=0}^{\infty}(\xi_k-\eta_k)c_k(\tau)\bigg|^2\mathrm{d}\tau\int_0^\beta\bigg(1+\Big(\sum_{k=0}^{\infty}\xi_k c_k(\tau)\Big)^2+\Big(\sum_{k=0}^{\infty}\eta_k c_k(\tau)\Big)^2\bigg)\mathrm{d}\tau} \\ &\leqslant \sqrt{3}M_1\sqrt{\sum_{k=0}^{\infty}|\xi_k-\eta_k|^2\Big(\beta+\sum_{k=0}^{\infty}|\xi_k|^2+\sum_{k=0}^{\infty}|\eta_k|^2\Big)}.\end{aligned}$$

Hence we obtain the desired result. □

Lemma A.3 Let $\pi_N(\xi)$ be the probability density function in $\mathbb{R}^{dN}$ defined in (3.4), i.e.,

$$\pi_N(\xi) = \frac{1}{Z_N}\exp\bigg(-\frac{1}{2}\sum_{k=0}^{N-1}(\omega_k^2+a^2)|\xi_k|^2 - \int_0^\beta V^a(x_N(\tau))\mathrm{d}\tau\bigg),$$

where Z_N is the normalization constant such that $\int_{\mathbb{R}^{dN}} \pi_N(\xi)\mathrm{d}\xi = 1$. Then

$$\int_{\mathbb{R}^{dN}} |\xi_k|^2 \pi_N(\xi)\mathrm{d}\xi \leqslant \frac{d}{\omega_k^2 + a^2} \exp\left(\frac{5}{2}\beta M_1 + C_0 M_1\right), \quad k = 0, 1, \cdots, N-1, \tag{A.8}$$

where the constant $C_0 = \frac{d\beta}{2a} \coth \frac{a\beta}{2}$. As a consequence, the second moment of $\pi_N(\xi)$ satisfies

$$\int_{\mathbb{R}^{dN}} \sum_{k=0}^{N-1} |\xi_k|^2 \pi_N(\xi)\mathrm{d}\xi \leqslant C_0 \exp\left(\frac{5}{2}\beta M_1 + C_0 M_1\right). \tag{A.9}$$

Proof We write the second moment of ξ_k in the distribution $\pi_N(\xi)$ by

$$\int_{\mathbb{R}^{dN}} |\xi_k|^2 \pi_N(\xi)\mathrm{d}\xi = \int_{\mathbb{R}^{dN}} |\xi_k|^2 \times \nu_N(\xi) \times \frac{\pi_N(\xi)}{\nu_N(\xi)}\mathrm{d}\xi, \quad k = 0, 1, \cdots, N-1, \tag{A.10}$$

where $\nu_N(\xi)$ is the Gaussian distributions of the normal mode coordinates $\{\xi_k\}_{k=0}^{N-1}$,

$$\nu_N(\xi) = \frac{1}{C_N} \exp\left(-\frac{1}{2}\sum_{k=0}^{N-1} (\omega_k^2 + a^2)|\xi_k|^2\right), \quad C_N = \int_{\mathbb{R}^{dN}} \exp\left(-\frac{1}{2}\sum_{k=0}^{N-1} (\omega_k^2 + a^2)|\xi_k|^2\right)\mathrm{d}\xi. \tag{A.11}$$

Since the potential function $V^a(q) \geqslant -M_1$ by Assumption (i), we have

$$\frac{\pi_N(\xi)}{\nu_N(\xi)} = \frac{C_N}{Z_N} \exp\left(-\int_0^\beta V^a(x_N(\tau))\mathrm{d}\tau\right) \leqslant \frac{C_N}{Z_N} \exp(\beta M_1). \tag{A.12}$$

Note that from (3.11) we have

$$V^a(q) \leqslant \frac{3}{2}M_1 + M_1|q|^2, \tag{A.13}$$

hence Z_N/C_N satisfies

$$\begin{aligned}
\frac{Z_N}{C_N} &= \int_{\mathbb{R}^{dN}} \exp\left(-\int_0^\beta V^a(x_N(\tau))\mathrm{d}\tau\right)\nu_N(\xi)\mathrm{d}\xi \\
&\geqslant \int_{\mathbb{R}^{dN}} \exp\left(-\int_0^\beta \left(\frac{3}{2}M_1 + M_1|x_N(\tau)|^2\right)\mathrm{d}\tau\right)\nu_N(\xi)\mathrm{d}\xi \\
&= \int_{\mathbb{R}^{dN}} \exp\left(-\frac{3}{2}\beta M_1 - M_1 \sum_{k=0}^{N-1} |\xi_k|^2\right)\nu_N(\xi)\mathrm{d}\xi \\
&= \exp\left(-\frac{3}{2}\beta M_1\right) \prod_{k=0}^{N-1} \left(\frac{\omega_k^2 + a^2}{2\pi}\right)^{\frac{d}{2}} \int_{\mathbb{R}^d} \exp\left(-\frac{2M_1 + \omega_k^2 + a^2}{2}|\xi_k|^2\right)\mathrm{d}\xi_k \\
&= \exp\left(-\frac{3}{2}\beta M_1\right) \prod_{k=0}^{N-1} \left(\frac{\omega_k^2 + a^2}{2M_1 + \omega_k^2 + a^2}\right)^{\frac{d}{2}}.
\end{aligned} \tag{A.14}$$

Now from Lemma A.1 we observe the equality

$$\sum_{k\in\mathbb{Z}} \frac{1}{k^2 + a^2} = \frac{\pi \coth(a\pi)}{a} \Longrightarrow \sum_{k=0}^{\infty} \frac{1}{\omega_k^2 + a^2} = \sum_{k\in\mathbb{Z}} \frac{1}{(\frac{2\pi k}{\beta})^2 + a^2} = \frac{C_0}{d}. \tag{A.15}$$

Then from the equality (A.15) we have

$$\prod_{k=0}^{N-1}\left(1+\frac{2M_1}{\omega_k^2+a^2}\right)\leqslant\prod_{k=0}^{N-1}\exp\left(\frac{2M_1}{\omega_k^2+a^2}\right)=\exp\left(\frac{2C_0}{d}M_1\right). \tag{A.16}$$

Then (A.14) yields

$$\frac{C_N}{Z_N}\leqslant\exp\left(\frac{3}{2}\beta M_1\right)\prod_{k=0}^{N-1}\left(1+\frac{2M_1}{\omega_k^2+a^2}\right)^{\frac{d}{2}}\leqslant\exp\left(\frac{3}{2}\beta M_1+C_0M_1\right), \tag{A.17}$$

and thus (A.12) implies

$$\frac{\pi_N(\xi)}{\nu_N(\xi)}\leqslant\exp\left(\frac{5}{2}\beta M_1+C_0M_1\right). \tag{A.18}$$

Using the upper bound estimate of $\pi_N(\xi)/\nu_N(\xi)$ in (A.18), the equality (A.10) gives

$$\int_{\mathbb{R}^{dN}}|\xi_k|^2\pi_N(\xi)\mathrm{d}\xi\leqslant\exp\left(\frac{5}{2}\beta M_1+C_0M_1\right)\int_{\mathbb{R}^{dN}}|\xi_k|^2\nu_N(\xi)\mathrm{d}\xi=\frac{d}{\omega_k^2+a^2}\exp\left(\frac{5}{2}\beta M_1+C_0M_1\right), \tag{A.19}$$

and thus (A.8) holds true. Using the equality (A.15), summation of (A.8) over k yields

$$\int_{\mathbb{R}^{dN}}\sum_{k=0}^{N-1}|\xi_k|^2\pi_N(\xi)\leqslant\sum_{k=0}^{\infty}\frac{d}{\omega_k^2+a^2}\exp\left(\frac{5}{2}\beta M_1+C_0M_1\right)=C_0\exp\left(\frac{5}{2}\beta M_1+C_0M_1\right),$$

which completes the proof. □

Corollary 3.1 Estimate the constant C_1.

Proof By Theorem 3.1, we have exponential decay of the relative entropy

$$H(\nu_NP_t|\pi_N)\leqslant\exp(-2\lambda_1t)H(\nu_N|\pi_N), \tag{A.20}$$

hence we only need to estimate the relative entropy $H(\nu_N|\pi_N)$ at the initial state, which characterizes the difference between the initial distribution $\nu_N(\xi)$ and the invariant distribution $\pi_N(\xi)$. The probability density of $\nu_N(\xi)$ in $\mathbb{R}^{dN}$ can be written as

$$\nu_N(\xi)=\frac{1}{C_N}\exp\left(-\frac{1}{2}\sum_{k=0}^{N-1}(\omega_k^2+a^2)|\xi_k|^2\right),\quad C_N=\int_{\mathbb{R}^{dN}}\exp\left(-\frac{1}{2}\sum_{k=0}^{N-1}(\omega_k^2+a^2)|\xi_k|^2\right)\mathrm{d}\xi, \tag{A.21}$$

hence the relative density $\nu_N(\xi)/\pi_N(\xi)$ is given by

$$\frac{\nu_N(\xi)}{\pi_N(\xi)}=\frac{Z_N}{C_N}\exp\left(\int_0^\beta V^a(x_N(\tau))\mathrm{d}\tau\right),\qquad\forall\xi\in\mathbb{R}^{dN}. \tag{A.22}$$

Since $V^a(q)\geqslant-M_1$ by Assumption (i), Z_N/C_N satisfies

$$\frac{Z_N}{C_N}=\int_{\mathbb{R}^{dN}}\exp\left(-\int_0^\beta V^a(x_N(\tau))\mathrm{d}\tau\right)\nu_N(\xi)\mathrm{d}\xi\leqslant\exp(\beta M_1). \tag{A.23}$$

Using the upper bound of $V^a(x(\tau))$ in (3.11), the relative entropy $H(\nu_N|\pi_N)$ is bounded by

$$
\begin{aligned}
H(\nu_N|\pi_N) &= \int_{\mathbb{R}^{dN}} \pi_N(\xi)\log\frac{\nu_N(\xi)}{\pi_N(\xi)}\mathrm{d}\xi \\
&\leqslant \log\frac{Z_N}{C_N} + \int_{\mathbb{R}^{dN}}\left(\int_0^\beta V^a(x_N(\tau))\mathrm{d}\tau\right)\pi_N(\xi)\mathrm{d}\xi \\
&\leqslant \beta M_1 + \int_{\mathbb{R}^{dN}}\int_0^\beta\left(\frac{3}{2}M_1 + M_1|x_N(\tau)|^2\right)\pi_N(\xi)\mathrm{d}\xi \\
&= \frac{5}{2}\beta M_1 + M_1\int_{\mathbb{R}^{dN}}\int_0^\beta\left|\sum_{k=0}^{N-1}\xi_k c_k(\tau)\right|^2\pi_N(\xi)\mathrm{d}\xi \\
&= \frac{5}{2}\beta M_1 + M_1\int_{\mathbb{R}^{dN}}\sum_{k=0}^{N-1}|\xi_k|^2\pi_N(\xi)\mathrm{d}\xi.
\end{aligned}
$$

Using the second moment estimate of $\pi_N(\xi)$ in Lemma A.3, we have

$$
H(\nu_N|\pi_N) \leqslant \frac{5}{2}\beta M_1 + C_0 M_1\exp\left(\frac{5}{2}\beta M_1 + C_0 M_1\right) = C_1. \tag{A.24}
$$

Hence $H(\nu_N P_t|\pi_N) \leqslant C_1\exp(-2\lambda_1 t)$ for any $t \geqslant 0$. □

Corollary 3.2 Estimate the constant C_2 in (3.89).

Proof For the initial distribution $\sigma_N(\xi,\eta)$, it is easy to deduce the corresponding relative density

$$
f(\xi,\eta) = \frac{\sigma_N(\xi,\eta)}{\mu_N(\xi,\eta)} = \frac{\nu_N(\xi)}{\pi_N(\xi)} \tag{A.25}
$$

actually does not depend on η, where $\nu_N(\xi)$ is the Gaussian distribution defined in (A.21), and $\pi_N(\xi)$ is the invariant distribution (3.4) of the truncated CL-PIMD. In Corollary 3.1, it has been proved that the relative entropy satisfies

$$
\mathrm{Ent}_{\mu_N}(f) = H(\sigma_N|\mu_N) = H(\nu_N|\pi_N) \leqslant C_1, \tag{A.26}
$$

hence $W(f)$ is estimated by

$$
\begin{aligned}
W(f) &= \left(\frac{M_3^2}{a^4}+1\right)\mathrm{Ent}_{\mu_N}(f) + \sum_{k=0}^{N-1}\frac{1}{\omega_k^2+a^2}\int_{\mathbb{R}^{dN}}\frac{|\nabla_{\xi_k}f|^2}{f}\mathrm{d}\pi_N \\
&\leqslant \left(\frac{M_3^2}{a^4}+1\right)C_1 + \sum_{k=0}^{N-1}\frac{1}{\omega_k^2+a^2}\int_{\mathbb{R}^{dN}}\frac{|\nabla_{\xi_k}f|^2}{f}\mathrm{d}\pi_N.
\end{aligned} \tag{A.27}
$$

Note that $f(\xi,\eta) = \nu_N(\xi)/\pi_N(\xi)$ can also be written as in the expression of (A.22),

$$
f(\xi,\eta) = \frac{Z_N}{C_N}\exp\left(\int_0^\beta V^a(x_N(\tau))\mathrm{d}\tau\right), \quad \xi,\eta\in\mathbb{R}^{dN}, \tag{A.28}
$$

hence taking the gradient with respect to ξ_k yields

$$\nabla_{\xi_k} f = f(\xi,\eta)\int_0^\beta \nabla V^a(x_N(\tau))c_k(\tau)\mathrm{d}\tau \Longrightarrow \frac{\nabla_{\xi_k} f}{f} = \int_0^\beta \nabla V^a(x_N(\tau))c_k(\tau)\mathrm{d}\tau. \tag{A.29}$$

Using the fact that the basis function $|c_k(\tau)| \leqslant \sqrt{2/\beta}$ for each k, from Assumption (i) we have

$$\left|\frac{\nabla_{\xi_k} f}{f}\right| \leqslant \sqrt{\frac{2}{\beta}}\int_0^\beta |\nabla V^a(x_N(\tau))|\mathrm{d}\tau \leqslant \sqrt{2\beta}M_1 + \sqrt{\frac{2}{\beta}}M_1\int_0^\beta |x_N(\tau)|\mathrm{d}\tau, \tag{A.30}$$

and thus using the Cauchy's inequality,

$$\begin{aligned}\left|\frac{\nabla_{\xi_k} f}{f}\right|^2 &\leqslant 4\beta M_1^2 + \frac{4M_1^2}{\beta}\left(\int_0^\beta |x_N(\tau)|\mathrm{d}\tau\right)^2 \\ &\leqslant 4\beta M_1^2 + 4M_1^2\int_0^\beta |x_N(\tau)|^2\mathrm{d}\tau = 4\beta M_1^2 + 4M_1^2\sum_{k=0}^{N-1}|\xi_k|^2.\end{aligned}$$

Hence the equality (A.15) implies for each $k = 0, 1, \cdots, N-1$,

$$\begin{aligned}\int_{\mathbb{R}^{dN}} \frac{|\nabla_{\xi_k} f|^2}{f}\mathrm{d}\pi_N = \int_{\mathbb{R}^{dN}}\left|\frac{\nabla_{\xi_k} f}{f}\right|^2\mathrm{d}\nu_N &\leqslant 4\beta M_1^2 + 4M_1^2\int_{\mathbb{R}^{dN}}\sum_{k=0}^{N-1}|\xi_k|^2\mathrm{d}\nu_N \\ &\leqslant 4\beta M_1^2 + 4M_1^2\sum_{k=0}^{\infty}\frac{d}{\omega_k^2+a^2} = 4(\beta + C_0)M_1^2.\end{aligned}$$

Then from (A.27) we have

$$W(f) \leqslant \left(\frac{M_3^2}{a^4}+1\right)C_1 + 4(\beta+C_0)M_1^2\sum_{k=0}^{N-1}\frac{1}{\omega_k^2+a^2} \leqslant \left(\frac{M_3^2}{a^4}+1\right)C_1 + \frac{4}{d}(\beta+C_0)C_0M_1^2. \tag{A.31}$$

Hence the constant C_2 is estimated as

$$C_2 = \frac{1}{M_3^2+1}W(f) \leqslant C_1 + \frac{4a^4(\beta+C_0)C_0}{d(M_3^2+a^4)}M_1^2, \tag{A.32}$$

which completes the proof. □

B Additional proofs for Section 3.4

Lemma B.1 For given positive integers M, N with $M > N$, and the real number $t \in [0,1]$, let $F(t,\xi)$ be the probability density function of $\xi = \{\xi_k\}_{k=0}^{M-1}$ defined in (B.25), i.e.,

$$F(t,\xi) = \frac{1}{Z_{M,N}}\exp\left(-\frac{1}{2}\sum_{k=0}^{M-1}(\omega_k^2+a^2)|\xi_k|^2 - t\int_0^\beta V^a\left(\sum_{k=0}^{M-1}\xi_k c_k(\tau)\right)\mathrm{d}\tau - (1-t)\int_0^\beta V^a\left(\sum_{k=0}^{N-1}\xi_k c_k(\tau)\right)\mathrm{d}\tau\right),$$

where $Z_{M,N}$ is the normalization constant such that $\int_{\mathbb{R}^{dM}} F(t,\xi)\mathrm{d}\xi = 1$. Then

$$\int_{\mathbb{R}^{dM}} |\xi_k|^2 F(t,\xi)\mathrm{d}\xi \leqslant \frac{d}{\omega_k^2 + a^2} \exp\bigg(\frac{5}{2}\beta M_1 + C_0 M_1\bigg), \qquad k = 0,1,\cdots,M-1, \tag{B.1}$$

where the constant $C_0 = \frac{d\beta}{2a}\coth\frac{a\beta}{2}$. As a consequence, the second moment of $F(t,\xi)$ satisfies

$$\int_{\mathbb{R}^{dM}} \sum_{k=0}^{M} |\xi_k|^2 F(t,\xi)\mathrm{d}\xi \leqslant C_0 \exp\bigg(\frac{5}{2}\beta M_1 + C_0 M_1\bigg) \tag{B.2}$$

Note that the case $t = 1$ recovers the result in Lemma A.3.

Proof The proof is similar to Lemma A.3. The second moment of ξ_k in the distribution $F(t,\xi)$ is

$$\int_{\mathbb{R}^{dM}} |\xi_k|^2 F(t,\xi)\mathrm{d}\xi = \int_{\mathbb{R}^{dM}} |\xi_k|^2 \times \nu_M(\xi) \times \frac{F(t,\xi)}{\pi_M(\xi)}\mathrm{d}\xi, \tag{B.3}$$

where $\nu_M(\xi)$ is the Gaussian distribution and C_M is the normalization constant,

$$\nu_M(\xi) = \frac{1}{C_M}\exp\bigg(-\frac{1}{2}\sum_{k=0}^{M-1}(\omega_k^2 + a^2)|\xi_k|^2\bigg), \;\; C_M = \int_{\mathbb{R}^{dM}} \exp\bigg(-\frac{1}{2}\sum_{k=0}^{M-1}(\omega_k^2 + a^2)|\xi_k|^2\bigg)\mathrm{d}\xi. \tag{B.4}$$

Then the relative density function $F(t,\xi)/\nu_M(\xi)$ in (B.3) is given by

$$\frac{F(t,\xi)}{\nu_M(\xi)} = \frac{C_M}{Z_{M,N}}\exp\bigg(-t\int_0^\beta V^a\bigg(\sum_{k=0}^{M-1}\xi_k c_k(\tau)\bigg)\mathrm{d}\tau - (1-t)\int_0^\beta V^a\bigg(\sum_{k=0}^{N-1}\xi_k c_k(\tau)\bigg)\mathrm{d}\tau\bigg), \tag{B.5}$$

and using the same arguments in Lemma A.3, we obtain

$$\frac{C_M}{Z_{M,N}} \leqslant \exp\bigg(\frac{3}{2}\beta M_1 + C_0 M_1\bigg), \tag{B.6}$$

and from (B.5) and $V^a(q) \geqslant -M_1$ by Assumption (i) we have

$$\frac{F(t,\xi)}{\nu_M(\xi)} \leqslant \exp\bigg(\frac{5}{2}\beta M_1 + C_0 M_1\bigg), \tag{B.7}$$

which is the counterpart of (A.18). The rest part of the proof is exactly the same as Lemma A.3.□

Lemma 3.1 Estimate the difference $|\langle O(\hat{q})\rangle_{\beta,N} - \langle O(\hat{q})\rangle_{\beta,M}|$.

Proof The proof is accomplished in several steps.

1. Expressions of probability distributions $\pi_N(\xi)$ and $\pi_M(\xi)$

For any positive integers M, N with $M > N$, define the loop average of the observable by

$$\bar{O}_N(\xi) = \frac{1}{\beta}\int_0^\beta O\bigg(\sum_{k=0}^{N-1}\xi_k c_k(\tau)\bigg)\mathrm{d}\tau, \qquad \bar{O}_M(\xi) = \frac{1}{\beta}\int_0^\beta O\bigg(\sum_{k=0}^{M-1}\xi_k c_k(\tau)\bigg)\mathrm{d}\tau. \tag{B.8}$$

Introduce the probability distributions in $\mathbb{R}^{dM}$ by

$$\bar{\pi}_N(\xi) = \frac{\exp(-U_N(\xi))}{\displaystyle\int_{\mathbb{R}^{dM}} \exp(-U_N(\xi))\mathrm{d}\xi}, \quad \bar{\pi}_M(\xi) = \frac{\exp(-U_M(\xi))}{\displaystyle\int_{\mathbb{R}^{dM}} \exp(-U_M(\xi))\mathrm{d}\xi}, \tag{B.9}$$

where both $U_N(\xi)$ and $U_M(\xi)$ are potential functions in $\mathbb{R}^{dM}$ defined by

$$U_N(\xi) = \frac{1}{2}\sum_{k=0}^{M-1}(\omega_k^2 + a^2)|\xi_k|^2 + \int_0^\beta V^a\bigg(\sum_{k=0}^{N-1}\xi_k c_k(\tau)\bigg)\mathrm{d}\tau, \tag{B.10}$$

$$U_M(\xi) = \frac{1}{2}\sum_{k=0}^{M-1}(\omega_k^2 + a^2)|\xi_k|^2 + \int_0^\beta V^a\bigg(\sum_{k=0}^{M-1}\xi_k c_k(\tau)\bigg)\mathrm{d}\tau. \tag{B.11}$$

Note that the difference between $U_N(\xi)$ and $U_M(\xi)$ only exists in the argument of the potential function $V^a(q)$. In this way, $\bar{\pi}_M(\xi)$ is exactly the invariant distribution $\pi_M(\xi)$ defined in (3.4), and $\bar{\pi}_N(\xi)$ is the product of $\pi_N(\xi)$ and the Gaussian distribution of $\{\xi_k\}_{k=N}^{M-1}$. Now the observable difference is then given by

$$\langle O(\hat{q})\rangle_{\beta,N} - \langle O(\hat{q})\rangle_{\beta,M} = \int_{\mathbb{R}^{dM}} \bar{O}_N(\xi)\bar{\pi}_N(\xi)\mathrm{d}\xi - \int_{\mathbb{R}^{dM}} \bar{O}_M(\xi)\bar{\pi}_M(\xi)\mathrm{d}\xi. \tag{B.12}$$

In order to estimate the observable difference, we write (B.12) as $I_1 + I_2$, where

$$I_1 = \int_{\mathbb{R}^{dM}} \big(\bar{O}_N(\xi) - \bar{O}_M(\xi)\big)\bar{\pi}_N(\xi)\mathrm{d}\xi, \quad I_2 = \int_{\mathbb{R}^{dM}} \bar{O}_M(\xi)\big(\bar{\pi}_N(\xi) - \bar{\pi}_M(\xi)\big). \tag{B.13}$$

Next we estimate I_1 and I_2 respectively.

2. Estimate I_1: difference in observable

By direct calculation, $\bar{O}_M(\xi) - \bar{O}_N(\xi)$ given by

$$\bar{O}_M(\xi) - \bar{O}_N(\xi) = \int_{\mathbb{R}^{dM}} \left\{\frac{1}{\beta}\int_0^\beta \left[O\bigg(\sum_{k=0}^{M-1}\xi_k c_k(\tau)\bigg) - O\bigg(\sum_{k=0}^{N-1}\xi_k c_k(\tau)\bigg)\right]\mathrm{d}\tau\right\}\bar{\pi}_N(\xi)\mathrm{d}\xi. \tag{B.14}$$

For any $q_1, q_2 \in \mathbb{R}^d$, define the function $f(\theta) = O(q_1 + \theta q_2)$ for $\theta \in [0,1]$, then the Taylor expansion of $f(\theta)$ at $\theta = 0$ yields

$$f(1) = f(0) + f'(0) + \int_0^1 f''(\theta)(1-\theta)\mathrm{d}\theta. \tag{B.15}$$

Note that we always have

$$f''(\theta) = q_2 \cdot \nabla^2 O(q_1 + \theta q_2)\cdot q_2 \Longrightarrow |f''(\theta)| \leqslant M_4|q_2|^2, \tag{B.16}$$

hence the equality (B.15) gives

$$\big|O(q_1+q_2) - O(q_1) - q_2\cdot\nabla O(q_1)\big| \leqslant M_4|q_2|^2 \int_0^1 (1-\theta)\mathrm{d}\theta = \frac{M_4}{2}|q_2|^2, \tag{B.17}$$

In the inequality (B.17) by choosing

$$q_1 = \sum_{k=0}^{N-1} \xi_k c_k(\tau), \qquad q_2 = \sum_{k=N}^{M-1} \xi_k c_k(\tau), \tag{B.18}$$

we have

$$\left| O\left(\sum_{k=0}^{M-1} \xi_k c_k(\tau)\right) - O\left(\sum_{k=0}^{N-1} \xi_k c_k(\tau)\right) - \sum_{k=N}^{M-1} \xi_k c_k(\tau) \cdot \nabla O\left(\sum_{k=0}^{N-1} \xi_k c_k(\tau)\right)\right| \leqslant \frac{M_4}{2}\left|\sum_{k=N}^{M-1} \xi_k c_k(\tau)\right|^2. \tag{B.19}$$

Since $\{c_k(\tau)\}_{k=0}^{M-1}$ is orthogonal in $L^2([0,\beta];\mathbb{R})$, integrating (B.19) in the interval $[0,\beta]$ gives

$$\frac{1}{\beta}\int_0^\beta \left| O\left(\sum_{k=0}^{M-1} \xi_k c_k(\tau)\right) - O\left(\sum_{k=0}^{N-1} \xi_k c_k(\tau)\right) - \sum_{k=N}^{M-1} \xi_k c_k(\tau) \cdot \nabla O\left(\sum_{k=0}^{N-1} \xi_k c_k(\tau)\right)\right| \mathrm{d}\tau \leqslant \frac{M_4}{2\beta}\sum_{k=N}^{M-1} |\xi_k|^2. \tag{B.20}$$

Note that $\xi \sim \bar{\pi}_N(\xi)$ implies

$$\xi_k \sim \mathcal{N}\left(0, \frac{I_d}{\omega_k^2 + a^2}\right), \qquad k = N, \cdots, M-1, \tag{B.21}$$

and $\omega_k \geqslant k\pi/\beta$ for any integer $k \geqslant 0$, we have

$$\int_{\mathbb{R}^{dM}} \sum_{k=N}^{M-1} |\xi_k|^2 \bar{\pi}_N(\xi) = \sum_{k=N}^{M-1} \frac{d}{\omega_k^2 + a^2} \leqslant \left(\frac{\beta}{\pi}\right)^2 \sum_{k=N}^{M-1} \frac{d}{k^2} \leqslant \frac{d\beta^2}{\pi^2(N-1)} \leqslant \frac{d\beta^2}{4N}. \tag{B.22}$$

Hence integrating (B.20) with the distribution $\bar{\pi}_N(\xi)$ in $\mathbb{R}^{dM}$ gives

$$\left| \int_{\mathbb{R}^{dM}} \left\{ \frac{1}{\beta}\int_0^\beta \left[O\left(\sum_{k=0}^{M-1} \xi_k c_k(\tau)\right) - O\left(\sum_{k=0}^{N-1} \xi_k c_k(\tau)\right)\right] \mathrm{d}\tau \right\} \bar{\pi}_N(\xi) \mathrm{d}\xi \right| \leqslant \frac{M_4}{2\beta} \times \frac{d\beta^2}{4N} = \frac{d\beta M_4}{8N}, \tag{B.23}$$

which implies $|I_1|$ has an upper bound $d\beta M_4/8N$.

3. Estimate I_2: difference in distribution

Since $\bar{O}_M(\xi)$ is uniformly bounded by M_4, $|I_2|$ is simply bounded by

$$|I_2| \leqslant M_4 \int_{\mathbb{R}^{dM}} |\bar{\pi}_N(\xi) - \bar{\pi}_M(\xi)| \mathrm{d}\xi, \tag{B.24}$$

which corresponds to the total variation between $\bar{\pi}_N(\xi)$ and $\bar{\pi}_M(\xi)$. To estimate $|I_2|$, define the parameter-dependent probability density function

$$F(t,\xi) = \frac{\exp\big(-tU_M(\xi) - (1-t)U_N(\xi)\big)}{\int_{\mathbb{R}^{dN}} \exp\big(-tU_M(\eta) - (1-t)U_N(\eta)\big)\mathrm{d}\eta}, \tag{B.25}$$

Note that the variable $\eta \in \mathbb{R}^{dN}$ here is to distinguish from $\xi \in \mathbb{R}^{dN}$, and does not refer to the auxiliary velocity variable to construct the underdamped Langevin dynamics (3.7). In other words, all the calculations here on involve the static properties of the invariant distributions.

It is easy to see $F(0,\xi) = \bar{\pi}_N(\xi)$ and $F(1,\xi) = \bar{\pi}_M(\xi)$, and the derivative of $F(t,\xi)$ with respect to the parameter t is given by

$$\frac{\partial}{\partial t}F(t,\xi) = F(t,\xi)\left[\big(U_N(\xi) - U_M(\xi)\big) - \int_{\mathbb{R}^{dM}} \big(U_N(\eta) - U_M(\eta)\big)F(t,\eta)\mathrm{d}\eta.\right]. \tag{B.26}$$

Applying Lemma A.2 on the two sets of normal mode coordinates

$$(\xi_0, \xi_1 \cdots, \xi_{N-1}, \underbrace{0, \cdots, 0}_{M-N}) \text{ and } (\xi_0, \xi_1, \cdots, \xi_{N-1}, \xi_N, \cdots, \xi_{M-1}), \tag{B.27}$$

we have the inequality

$$|U_M(\xi) - U_N(\xi)| \leqslant \sqrt{3}M_1 \sqrt{\sum_{k=N}^{M-1} |\xi_k|^2 \left(\beta + 2\sum_{k=0}^{M-1} |\xi_k|^2\right)}. \tag{B.28}$$

Taking the expectation in the distribution $F(t,\xi)$ in (B.28) gives

$$\begin{aligned}
&\int_{\mathbb{R}^{dM}} |U_M(\xi) - U_N(\xi)|F(t,\xi)\mathrm{d}\xi \\
\leqslant{}& \sqrt{3}M_1 \int_{\mathbb{R}^{dM}} \sqrt{\sum_{k=N}^{M-1} |\xi_k|^2 \left(\beta + 2\sum_{k=0}^{M-1} |\xi_k|^2\right)} F(t,\xi)\mathrm{d}\xi \\
\leqslant{}& \sqrt{3}M_1 \sqrt{\int_{\mathbb{R}^{dM}} \sum_{k=N}^{M-1} |\xi_k|^2 F(t,\xi)\mathrm{d}\xi \times \left(\beta + 2\int_{\mathbb{R}^{dM}} \sum_{k=0}^{M-1} |\xi_k|^2 F(t,\xi)\mathrm{d}\xi\right)}.
\end{aligned}$$

Using the inequality

$$\sum_{k=N}^{M-1} \frac{d}{\omega_k^2 + a^2} \leqslant \frac{d\beta^2}{4N} \tag{B.29}$$

derived in (B.22) and the second moment estimate in Lemma B.1, we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^{dM}} |U_M(\xi) - U_N(\xi)|F(t,\xi)\mathrm{d}\xi \\
\leqslant{}& \sqrt{3}M_1 \sqrt{\sum_{k=N}^{M-1} \frac{d}{\omega_k^2 + a^2} \exp\left(\frac{5}{2}\beta M_1 + C_0 M_1\right) \times \left(\beta + 2C_0 \exp\left(\frac{5}{2}\beta M_1 + C_0 M_1\right)\right)} \\
\leqslant{}& \sqrt{3}M_1 \exp\left(\frac{5}{2}\beta M_1 + C_0 M_1\right) \sqrt{\frac{d\beta^2}{4N} \times (\beta + 2C_0)} \\
\leqslant{}& \sqrt{d(\beta + 2C_0)}\beta M_1 \exp\left(\frac{5}{2}\beta M_1 + C_0 M_1\right) \frac{1}{\sqrt{N}}.
\end{aligned} \tag{B.30}$$

From (B.26) and (B.30), we have

$$\left|\frac{\partial F}{\partial t}\right| \leqslant F(t,\xi)\big|U_M(\xi) - U_N(\xi)\big| + \sqrt{d(\beta + 2C_0)}\beta M_1 \exp\left(\frac{5}{2}\beta M_1 + C_0 M_1\right) \frac{F(t,\xi)}{\sqrt{N}}. \tag{B.31}$$

Integrating (B.31) over all the normal mode coordinates $\xi = \{\xi_k\}_{k=0}^{M-1}$ gives

$$\int_{\mathbb{R}^{dM}} \left| \frac{\partial F}{\partial t}(t,\xi) \right| \mathrm{d}\xi \leqslant 2\sqrt{d(\beta + 2C_0)}\beta M_1 \exp\left(\frac{5}{2}\beta M_1 + C_0 M_1\right)\frac{1}{\sqrt{N}}. \tag{B.32}$$

Finally, integrating (B.32) in $t \in [0,1]$ yields

$$\begin{aligned}\int_{\mathbb{R}^{dM}} |\bar{\pi}_N(\xi) - \bar{\pi}_M(\xi)| \mathrm{d}\xi &= \int_{\mathbb{R}^{dM}} |F(1,\xi) - F(0,\xi)| \mathrm{d}\xi \\ &\leqslant \int_{\mathbb{R}^{dM}} \left[\int_0^1 \left|\frac{\partial F}{\partial t}(t,\xi)\right| \mathrm{d}t\right] \mathrm{d}\xi \leqslant \max_{0\leqslant t\leqslant 1}\left\{\int_{\mathbb{R}^{dM}} \left|\frac{\partial F}{\partial t}(t,\xi)\right| \mathrm{d}\xi\right\} \\ &= 2\sqrt{d(\beta + 2C_0)}\beta M_1 \exp\left(\frac{5}{2}\beta M_1 + C_0 M_1\right)\frac{1}{\sqrt{N}}.\end{aligned}$$

Therefore we obtain

$$|I_2| \leqslant 2\sqrt{d(\beta + 2C_0)}\beta M_1 M_4 \exp\left(\frac{5}{2}\beta M_1 + C_0 M_1\right)\frac{1}{\sqrt{N}}. \tag{B.33}$$

4. Combining estimates of I_1 and I_2

Combining the estimates of $|I_1|$ and $|I_2|$, we obtain

$$|\langle O(\hat{q})\rangle_{\beta,N} - \langle O(\hat{q})\rangle_{\beta,M}| \leqslant \frac{d\beta M_4}{8N} + 2\sqrt{d(\beta + 2C_0)}\beta M_1 M_4 \exp\left(\frac{5}{2}\beta M_1 + C_0 M_1\right)\frac{1}{\sqrt{N}}. \tag{B.34}$$

□

Lemma 3.2 Identify the limit of $\langle O(\hat{q})\rangle_{\beta,N}$ as $N \to \infty$.

Proof Using Lemma A.2, we have the inequality

$$\begin{aligned}&\left|\exp\left(-\int_0^\beta V^a(x_N(\tau))\mathrm{d}\tau\right) - \exp\left(-\int_0^\beta V^a(x(\tau))\mathrm{d}\tau\right)\right| \\ &\leqslant \int_0^\beta \big|V^a(x_N(\tau)) - V^a(x(\tau))\big|\mathrm{d}\tau \leqslant \sqrt{3}M_1\sqrt{\sum_{k=N}^\infty |\xi_k|^2\bigg(\beta + 2\sum_{k=0}^\infty |\xi_k|^2\bigg)}.\end{aligned} \tag{B.35}$$

Using the Cauchy's inequality, (B.35) implies

$$\begin{aligned}&\mathbb{E}_\nu\left|\exp\left(-\int_0^\beta V^a(x_N(\tau))\mathrm{d}\tau\right) - \exp\left(-\int_0^\beta V^a(x(\tau))\mathrm{d}\tau\right)\right|^2 \\ &\leqslant 3M_1^2\,\mathbb{E}_\nu\bigg[\sum_{k=N}^\infty |\xi_k|^2\bigg(\beta + 2\sum_{k=0}^\infty |\xi_k|^2\bigg)\bigg] \\ &= 3M_1^2\,\mathbb{E}_\nu\bigg[\sum_{k=N}^\infty |\xi_k|^2\bigg(\beta + 2\sum_{k=0}^{N-1} |\xi_k|^2\bigg) + 2\sum_{k=N}^\infty |\xi_k|^4\bigg] \\ &= 3M_1^2\bigg[\sum_{k=N}^\infty \frac{d}{\omega_k^2 + a^2}\bigg(\beta + 2\sum_{k=0}^{N-1}\frac{d}{\omega_k^2 + a^2}\bigg) + 2\sum_{k=N}^\infty \frac{d(d+2)}{(\omega_k^2 + a^2)^4}\bigg].\end{aligned} \tag{B.36}$$

Hence from (B.36) we obtain

$$\lim_{N\to\infty} \mathbb{E}_\nu \left| \exp\left(- \int_0^\beta V^a(x_N(\tau))\mathrm{d}\tau \right) - \exp\left(- \int_0^\beta V^a(x(\tau))\mathrm{d}\tau \right) \right|^2 = 0, \tag{B.37}$$

which implies the random variable

$$\exp\left(- \int_0^\beta V^a(x_N(\tau))\mathrm{d}\tau \right) \xrightarrow{L^2(\nu)} \exp\left(- \int_0^\beta V^a(x(\tau))\mathrm{d}\tau \right), \quad \text{as } N \to \infty. \tag{B.38}$$

Since in Assumption (iv) the observable function $O(q)$ has bounded derivatives, we can apply the same arguments on $O(q)$ to conclude that

$$\frac{1}{\beta}\int_0^\beta O(x_N(\tau))\mathrm{d}\tau \xrightarrow{L^2(\nu)} \frac{1}{\beta}\int_0^\beta O(x(\tau))\mathrm{d}\tau, \quad \text{as } N \to \infty. \tag{B.39}$$

From (B.38) and (B.39) we obtain the desired result. □

C Generalized Γ calculus for degenerate diffusion processes

We briefly review the generalized Γ calculus developed in [27,28], which deals with the ergodicity of the Markov processes with degenerate diffusions, for example, the underdamped Langevin dynamics. The generalized Γ calculus is based on the classical Bakry–Émery theory [26].

Let $\{X_t\}_{t\geqslant 0}$ be a reversible Markov process in $\mathbb{R}^d$, and $(P_t)_{t\geqslant 0}$ be the Markov semigroup. Let L be the infinitesimal generator of $\{X_t\}_{t\geqslant 0}$, and π be the invariant distribution. Then L is self-adjoint in $L^2(\pi)$. In the classical Bakry–Émery theory, the carré du champ operator $\Gamma_1(f,g)$ and the iterated operator $\Gamma_2(f,g)$ are defined by

$$\Gamma_1(f,g) = \frac{1}{2}(L(fg) - gLf - fLg), \tag{C.1}$$

and

$$\Gamma_2(f,g) = \frac{1}{2}(L\Gamma_1(f,g) - \Gamma_1(f,Lg) - \Gamma_1(g,Lf)). \tag{C.2}$$

For notational convenience, we simply write $\Gamma_1(f) = \Gamma_1(f,f)$ and $\Gamma_2(f) = \Gamma_2(f,f)$. The curvature-dimension condition $CD(\rho,\infty)$ is known as the function inequality

$$\Gamma_2(f) \geqslant \rho\Gamma_1(f), \quad \text{for any smooth function } f. \tag{C.3}$$

For a given positive smooth function f in $\mathbb{R}^d$, define the relative entropy of f with respect to the invariant distribution π by

$$\mathrm{Ent}_\pi(f) = \int_{\mathbb{R}^d} f\log f\mathrm{d}\pi - \int_{\mathbb{R}^d} f\mathrm{d}\pi \log \int_{\mathbb{R}^d} f\mathrm{d}\pi. \tag{C.4}$$

If $\rho > 0$, then $CD(\rho,\infty)$ implies the log-Sobolev inequality (see Equation (5.7.1) of [26])

$$\mathrm{Ent}_\pi(f) \leqslant \frac{1}{2\rho}\int_{\mathbb{R}^d} \frac{\Gamma_1(f)}{f}\mathrm{d}\pi, \quad \text{for any positive smooth function } f. \tag{C.5}$$

and thus the exponential decay of the relative entropy (see Theorem 5.2.1 of [26])

$$\mathrm{Ent}_\pi(P_t f) \leqslant e^{-2\rho t}\,\mathrm{Ent}_\pi(f), \quad \forall t \geqslant 0. \tag{C.6}$$

Inspired from the operators Γ_1 and Γ_2, we define the generalized Γ operator as follows.

Definition C.1 Suppose f is a smooth function, and $\Phi(f)$ be a function of f and its derivatives. For a given stochastic process $\{X_t\}_{t\geqslant 0}$ with generator L, define the generalized Γ operator by

$$\Gamma_\Phi(f) = \frac{1}{2}(L\Phi(f) - \mathrm{d}\Phi(f)\cdot Lf), \tag{C.7}$$

where $\mathrm{d}\Phi(f)\cdot g$ for two smooth functions f, g are given by

$$\mathrm{d}\Phi(f)\cdot g = \lim_{s\to 0}\frac{\Phi(f+sg)-\Phi(f)}{s}. \tag{C.8}$$

The expression of $\Gamma_\Phi(f)$ can be conveniently obtained via the following result (Lemma 5 of [28]).

Lemma C.1 Suppose C_1, C_2 are two linear operators and $\Phi(f) = C_1 f \cdot C_2 f$, then

$$\Gamma_\Phi(f) = \Gamma_1(C_1 f, C_2 f) + \frac{1}{2}C_1 f\cdot [L, C_2]f + \frac{1}{2}[L, C_1]f\cdot C_2 f, \tag{C.9}$$

where $\Gamma(\cdot,\cdot)$ is the classical carré du champ operator.

In the following we assume the stochastic process $\{X_t\}_{t\geqslant 0}$ in $\mathbb{R}^d$ is solved by the SDE

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma \mathrm{d}B_t, \quad t \geqslant 0, \tag{C.10}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$ is the drift force, $\sigma \in \mathbb{R}^{d\times m}$ is a constant matrix, and $\{B_t\}_{t\geqslant 0}$ is the standard Brownian motion in $\mathbb{R}^m$. Then the generator of $\{X_t\}_{t\geqslant 0}$ is given by

$$Lf(x) = b(x)\cdot f(x) + \nabla\cdot(D\nabla f), \tag{C.11}$$

where $D = \frac{1}{2}\sigma\sigma^{\mathrm{T}} \in \mathbb{R}^{d\times d}$ is the constant diffusion matrix. For the generator L given in (C.11), we calculate the Γ operator $\Gamma_\Phi(f)$ with some classical choices of Φ.

Example C.1 If $\Phi(f) = |f|^2$, then we can take $C_1 = C_2 = 1$ in Lemma C.1 to obtain

$$\Gamma_\Phi(f) = \Gamma_1(f) = \frac{1}{2}\big[\nabla(D\nabla(f^2)) - 2f\nabla(D\nabla f)\big] = (\nabla f)^{\mathrm{T}} D\nabla f. \tag{C.12}$$

In particular, since $D \in \mathbb{R}^{d\times d}$ is positive semidefinite, we always have $\Gamma_1(f) \geqslant 0$.

Example C.2 If $\Phi(f) = f\log f$, then by direct calculation we have

$$\Gamma_\Phi(f) = \frac{1}{2}\Big[\nabla\cdot\big(D(\log f + 1)\nabla f\big) - (\log f + 1)Lf\Big] = \frac{(\nabla f)^{\mathrm{T}} D\nabla f}{2f}. \tag{C.13}$$

Example C.3 If $\Phi(f) = |Cf|^2$ for some linear operator C, then from Lemma C.1

$$\Gamma_\Phi(f) = \Gamma(Cf) + Cf\cdot[L, Cf]. \tag{C.14}$$

From Example C.1 we have $\Gamma_1(Cf) \geqslant 0$, hence we obtain the convenient estimate

$$\Gamma_\Phi(f) \geqslant Cf\cdot[L, Cf]. \tag{C.15}$$

Example C.4 If $\Phi(f) = |Cf|^2/f$ for some linear operator C, then

$$\Gamma_\Phi(f) \geqslant \frac{Cf \cdot [L, Cf]}{f}. \tag{C.16}$$

The proof below is given in Lemma 7 of [28].

Proof It is easy to verify for any smooth functions f, g, there holds

$$L(fg) = gLf + fLg + 2\Gamma_1(f, g). \tag{C.17}$$

By replacing $f \to |Cf|^2$ and $g \to 1/f$, we have

$$\begin{aligned} L\left(\frac{|Cf|^2}{f}\right) &= \frac{1}{f}L(|Cf|^2) + |Cf|^2 L\left(\frac{1}{f}\right) + 2\Gamma_1\left(|Cf|^2, \frac{1}{f}\right) \\ &= \frac{1}{f}L(|Cf|^2) + |Cf|^2\left(-\frac{Lf}{f^2} + \frac{2}{f^3}\Gamma_1(f)\right) + \frac{4Cf \cdot \Gamma_1(Cf, f)}{f^2}. \end{aligned} \tag{C.18}$$

Note that

$$\mathrm{d}\left(\frac{|Cf|^2}{f}\right) \cdot |Cf|^2 = \frac{\mathrm{d}(|Cf|^2) \cdot Lf}{f^2} - |Cf|^2 \frac{Lf}{f^2}. \tag{C.19}$$

Hence from (C.18)(C.19) and the definition of the generalized Γ operator in (C.1), we obtain

$$\Gamma_\Phi(f) \geqslant \frac{\Gamma_{|C\cdot|^2}(f)}{f} + \frac{1}{f^3}|Cf|^2\Gamma_1(f) + \frac{2Cf \cdot \Gamma_1(Cf, f)}{f^2}, \tag{C.20}$$

where $\Gamma_{|C\cdot|^2}$ is the generalized Γ operator induced by the function $|Cf|^2$. Since the matrix $D \in \mathbb{R}^{d\times d}$ is positive semidefinite, we have the Cauchy-Swarchz inequality

$$|\Gamma_1(f, Cf)|^2 \leqslant \Gamma_1(f)\Gamma_1(Cf), \quad \text{for any smooth function } f. \tag{C.21}$$

From (C.20) and (C.21) we obtain

$$\Gamma_\Phi(f) \geqslant \frac{\Gamma_1(Cf) + Cf \cdot [L, C]f}{f} + \frac{|Cf|^2\Gamma_1(f)}{f^3} - 2\sqrt{\frac{|Cf|^2\Gamma_1(f)}{f^3}}\sqrt{\frac{\Gamma_1(Cf)}{f}} \geqslant \frac{Cf \cdot [L, C]f}{f}. \tag{C.22}$$

Hence we obtain the desired result. $\square$

Now we establish the curvature-dimension conditions for using the generalized Γ operators. The following result relates the time derivative of $\Phi(f)$ with the operator $\Gamma_\Phi(f)$.

Lemma C.2 Given the constant $t > 0$, for any $s \in [0, t]$, we have the equality

$$\frac{\mathrm{d}}{\mathrm{d}s}\big[P_s\Phi(P_{t-s}f)(x)\big] = 2P_s\Gamma_\Phi(P_{t-s}f)(x). \tag{C.23}$$

As a consequence, for any $t \geqslant 0$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} \Phi(P_t f)\mathrm{d}\pi = -2\int_{\mathbb{R}^d} \Gamma_\Phi(P_t f)\mathrm{d}\pi. \tag{C.24}$$

Proof Using the chain rule, we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\big[P_s\Phi(P_{t-s}f)\big] &= LP_s\Phi(P_{t-s}f) + P_s\frac{\mathrm{d}}{\mathrm{d}s}\big[\Phi(P_{t-s}f)\big]\\ &= LP_s\Phi(P_{t-s}f) + P_s\lim_{r\to 0}\frac{\Phi(P_{t-s-r}f)-\Phi(P_{t-s}f)}{r}\\ &= LP_s\Phi(P_{t-s}f) - P_s\mathrm{d}\Phi(P_{t-s}f)\cdot LP_{t-s}f\\ &= P_s\big(L\Phi_{t-s}f - \mathrm{d}\Phi(P_{t-s}f)\cdot LP_{t-s}f\big)\\ &= 2P_s\Gamma_\Phi(P_{t-s}f).\end{aligned}$$

Integrating the equality over the distribution π, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}s}\int_{\mathbb{R}^d} P_s\Phi(P_{t-s}f)\mathrm{d}\pi = 2\int_{\mathbb{R}^d}\Gamma_\Phi(P_{t-s}f)\mathrm{d}\pi. \tag{C.25}$$

Since π is the invariant distribution, we obtain from (C.25)

$$\frac{\mathrm{d}}{\mathrm{d}s}\int_{\mathbb{R}^d}\Phi(P_{t-s}f)\mathrm{d}\pi = 2\int_{\mathbb{R}^d}\Gamma_\Phi(P_{t-s}f)\mathrm{d}\pi. \tag{C.26}$$

Replacing $t-s$ by s in (C.26), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}s}\int_{\mathbb{R}^d}\Phi(P_tf)\mathrm{d}\pi = -2\int_{\mathbb{R}^d}\Gamma_\Phi(P_tf)\mathrm{d}\pi, \tag{C.27}$$

which completes the proof. □

By choosing $\Phi(f) = f\log f$, then Lemma C.2 implies

$$\frac{\mathrm{d}}{\mathrm{d}t}\operatorname{Ent}_\pi(P_tf) = \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d}\Phi(P_tf)\mathrm{d}t = -\int_{\mathbb{R}^d}\frac{(\nabla f)^{\mathrm{T}}D\nabla f}{f}\mathrm{d}\pi = -\int_{\mathbb{R}^d}\frac{\Gamma(f)}{f}\mathrm{d}\pi. \tag{C.28}$$

If we have the log-Sobolev inequality (C.5), from (C.28) we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\operatorname{Ent}_\pi(P_tf) \leqslant -2\rho\operatorname{Ent}_\pi(f), \tag{C.29}$$

which implies $\operatorname{Ent}_\pi(P_tf) \leqslant e^{-2\rho t}\operatorname{Ent}_\pi(f)$, and we recover the result in (C.6).

Now we state the main theorem, which provides the curvature-dimension condition for degenerate diffusion processes.

Theorem C.1 Let $\{X_t\}_{t\geqslant 0}$ be an ergodic stochastic process with invariant distribution π. If for two functions $\Phi_1(f)$ and $\Phi_2(f)$, there holds

$$0 \leqslant \int_{\mathbb{R}^d}\Phi_1(f)\mathrm{d}\pi - \Phi_1\bigg(\int_{\mathbb{R}^d} f\mathrm{d}\pi\bigg) \leqslant c\int_{\mathbb{R}^d}\Phi_2(f)\mathrm{d}\pi, \quad \text{for any smooth function } f \tag{C.30}$$

and

$$\Gamma_{\Phi_2}(f) \geqslant \rho\Phi_2(f) - m\Gamma_{\Phi_1}(f), \quad \text{for any smooth function } f \tag{C.31}$$

for some $m > 0$, then for the function

$$W(f) = \beta\bigg(\int_{\mathbb{R}^d} \Phi_1(f)\mathrm{d}\pi - \Phi_1\bigg(\int_{\mathbb{R}^d} f\mathrm{d}\pi\bigg)\bigg) + \int_{\mathbb{R}^d} \Phi_2(f)\mathrm{d}\pi, \tag{C.32}$$

there holds

$$W(P_t f) \leqslant \exp\bigg(-\frac{2\rho t}{1+mc}\bigg)W(f), \quad \forall t \geqslant 0. \tag{C.33}$$

The proof below is given in Lemma 3 of [28].

Proof Using Lemma C.2 and (C.31), we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}W(P_t f) &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg[\beta\int_{\mathbb{R}^d} \Phi_1(P_t f)\mathrm{d}\pi + \int_{\mathbb{R}^d} \Phi_2(P_t f)\mathrm{d}\pi\bigg] \\ &= -2\int_{\mathbb{R}^d} \big(m\Gamma_{\Phi_1} + \Gamma_{\Phi_2}\big)(P_t f)\mathrm{d}\pi \leqslant -2\rho\int_{\mathbb{R}^d} \Phi_2(P_t f)\mathrm{d}\pi.\end{aligned} \tag{C.34}$$

Using (C.30) and the definiton of $W(f)$, we have

$$W(f) = m\bigg(\int_{\mathbb{R}^d} \Phi_1(f)\mathrm{d}\pi - \Phi_1\bigg(\int_{\mathbb{R}^d} f\mathrm{d}\pi\bigg)\bigg) + \int_{\mathbb{R}^d} \Phi_2(f)\mathrm{d}\pi \leqslant (1+mc)\int_{\mathbb{R}^d} \Phi_2(f)\mathrm{d}\pi. \tag{C.35}$$

Hence (C.34) implies

$$\frac{\mathrm{d}}{\mathrm{d}t}W(P_t f) \leqslant -\frac{2\rho}{1+mc}W(P_t f), \quad \forall t \geqslant 0, \tag{C.36}$$

yielding the desired result. □

As a heuristic example, we apply the generalized Γ calculus to study the ergodicity of the following linear underdamped Langevin dynamics. Let $(x_t, v_t) \in \mathbb{R}^d \times \mathbb{R}^d$ be evolved by

$$\begin{cases} \dot{x}_t = v_t, \\ \dot{v}_t = -x_t - v_t + \sqrt{\dfrac{2}{\beta}}\dot{B}_t, \end{cases} \tag{C.37}$$

where $\beta > 0$ is the inverse temperature, and B_t is the standard Brownian motion in $\mathbb{R}^d$. The generator of (C.37) is given by

$$L = v \cdot \nabla_x - (x+v) \cdot \nabla_v + \frac{1}{\beta}\Delta_v, \tag{C.38}$$

and the invariant distribution is

$$\mu(x,v) \propto \exp\bigg(-\frac{\beta}{2}(|x|^2 + |v|^2)\bigg). \tag{C.39}$$

A brief demonstration of the proof for the ergodicity of (C.37) is given as follows.

1. The log-Sobolev inequality for $\mu(x,v)$ is given by

$$\mathrm{Ent}_\mu(f) \leqslant \frac{1}{2}\int_{\mathbb{R}^d} \frac{|\nabla_x f|^2 + |\nabla_v f|^2}{f}\mathrm{d}\mu, \quad \text{for any positive function } f. \tag{C.40}$$

2. Inspired from Example 3 of [28], we choose the functions $\Phi_1(f)$ and $\Phi_2(f)$ by

$$\Phi_1(f) = f \log f, \quad \Phi_2(f) = \frac{|\nabla_v f - \nabla_x f|^2 + |\nabla_v f|^2}{f}. \tag{C.41}$$

Using Example C.2, $\Gamma_{\Phi_1}(f)$ is given by

$$\Gamma_{\Phi_1}(f) = \frac{1}{2\beta} \frac{|\nabla_v f|^2}{f}. \tag{C.42}$$

Using Example C.4, $\Gamma_{\Phi_2}(f)$ is estimated as

$$\Gamma_{\Phi_2}(f) \geqslant \frac{(\nabla_v - \nabla_x) f \cdot [L, (\nabla_v - \nabla_x)] f}{f} + \frac{\nabla_v f \cdot [L, \nabla_v] f}{f} = \frac{|\nabla_v f - \nabla_x f|^2}{f}. \tag{C.43}$$

3. Now we have the log-Sobolev inequality

$$0 \leqslant \int_{\mathbb{R}^d} \Phi_1(f) \mathrm{d}\mu - \Phi_1\left(\int_{\mathbb{R}^d} f \mathrm{d}\mu \right) \leqslant \frac{3}{2} \int_{\mathbb{R}^d} \Phi_2(f) \mathrm{d}\mu, \quad \text{for any smooth function } f \tag{C.44}$$

and the generalized curvature-dimension condition

$$\Gamma_{\Phi_2}(f) - \frac{1}{2}\Phi_2(f) \geqslant \frac{|\nabla_v f - \nabla_x f|^2 - |\nabla_v f|^2}{2f} \geqslant -\beta \Gamma_{\Phi_1}(f). \tag{C.45}$$

Then we can take the constants

$$c = \frac{3}{2}, \quad \rho = \frac{1}{2}, \quad m = \beta \tag{C.46}$$

in Theorem C.1 to obtain the geometric ergodicity of (C.37) in the sense of the relative entropy.

D Spectral method to compute the thermal average

When the spatial dimension $d = 1$, we present a benchmark spectral method to compute the quantum thermal average of $\langle O(\hat{q}) \rangle_\beta$ for given $\beta > 0$. The core principle of this method is the pseudo-spectral method [41] or the discrete variable representation (DVR) [42]. The method starts with the spectrum of the quantum harmonic oscillator

$$\hat{H}^a = -\frac{\Delta_d}{2} + \frac{a^2}{2} q^2, \tag{D.1}$$

whose eigenvalues and eigenfunctions are explicitly given by (see Section 1.5 of [24])

$$E_n = \left(n + \frac{1}{2} \right) a, \quad \psi_n(q) = \frac{1}{\sqrt{2^n n!}} \left(\frac{a}{\pi} \right)^{\frac{1}{4}} e^{-\frac{a}{2} q^2} H_n(aq), \quad n = 0, 1, 2, \cdots, \tag{D.2}$$

where $H_n(z)$ is the Hermite polynomial of order n

$$H_n(z) = (-1)^n e^{z^2} \frac{\mathrm{d}^n}{\mathrm{d}z^n} (e^{-z^2}), \quad n = 0, 1, 2, \cdots. \tag{D.3}$$

Also, $\{\psi_n(q)\}_{n=0}^{\infty}$ forms an orthonormal set of wavefunctions, hence we can write

$$\mathrm{Tr}[e^{-\beta\hat{H}}] = \sum_{n=0}^{\infty} \langle\psi_n|e^{-\beta\hat{H}}|\psi_n\rangle\,, \qquad \mathrm{Tr}[e^{-\beta\hat{H}}O(\hat{q})] = \sum_{n=0}^{\infty} \langle\psi_n|e^{-\beta\hat{H}}O(\hat{q})|\psi_n\rangle\,. \tag{D.4}$$

Define $\varphi(\tau,q) = e^{-\tau\hat{H}}\psi_n(q)$ for $\tau \geqslant 0$, then $\varphi(\tau,q)$ is solved by the parabolic equation

$$\dot{\varphi}(\tau,q) = -\hat{H}\varphi(\tau,q) = -\hat{H}^a\varphi(\tau,q) - V^a(q)\varphi(\tau,q) \tag{D.5}$$

with initial value $\varphi(0,q) = \psi_n(q)$. Since $\{\psi_n(q)\}_{n=0}^{\infty}$ is the basis function in $L^2(\mathbb{R})$, we can write

$$\varphi(\tau,q) = \sum_{n=0}^{\infty} c_n(\tau)\psi_n(q), \tag{D.6}$$

where the coefficients $\{c_n\}_{n=0}^{\infty}$ are compactly written as an infinite-dimensional vector $\boldsymbol{c}$. Note that

$$\|\varphi(\tau,\cdot)\|_{L^2(\mathbb{R})}^2 = \sum_{n=0}^{\infty} c_n(\tau)^2, \tag{D.7}$$

we conclude that the vector $\boldsymbol{c}$ lives in the Hilbert space l^2, the space of square summable real sequences (see Section 6.1 of [43]). Moreover, define the energy operator $\boldsymbol{E}$ in l^2 by

$$(\boldsymbol{E}\boldsymbol{c})_n = E_n c_n, \qquad n = 0,1,2,\cdots \tag{D.8}$$

the potential operator $\boldsymbol{V}^a$ in l^2 by

$$(\boldsymbol{V}^a\boldsymbol{c})_n = \sum_{m=0}^{\infty} \langle\psi_n|V^a|\psi_m\rangle\, c_m, \qquad n = 0,1,2,\cdots\,, \tag{D.9}$$

and the observable operator $\boldsymbol{O}$ in l^2 by

$$(\boldsymbol{O}\boldsymbol{c})_n = \sum_{m=0}^{\infty} \langle\psi_n|O|\psi_m\rangle\, c_m, \qquad n = 0,1,2,\cdots\,. \tag{D.10}$$

The elements $\langle\psi_n|V^a|\psi_m\rangle$ and $\langle\psi_n|O|\psi_m\rangle$ can be computed using the Gauss–Hermite quadrature. Then we can rewrite the parbolic equation (D.5) as

$$\dot{\boldsymbol{c}}(\tau) = -(\boldsymbol{E}+\boldsymbol{V}^a)\boldsymbol{c}(\tau) \Longrightarrow \boldsymbol{c}(\beta) = e^{-\beta(\boldsymbol{E}+\boldsymbol{V}^a)}\boldsymbol{e}_n, \tag{D.11}$$

where $\boldsymbol{e}_n$ is the unit vector in l^2 with the n-th entry being 1. Therefore the trace terms $\mathrm{Tr}[e^{-\beta\hat{H}}]$ and $\mathrm{Tr}[e^{-\beta\hat{H}}O(\hat{q})]$ can be computed by

$$\mathrm{Tr}[e^{-\beta\hat{H}}] = \mathrm{Tr}[e^{-\beta(\boldsymbol{E}+\boldsymbol{V}^a)}], \qquad \mathrm{Tr}[e^{-\beta\hat{H}}O(\hat{q})] = \mathrm{Tr}[e^{-\beta(\boldsymbol{E}+\boldsymbol{V}^a)}\boldsymbol{O}]. \tag{D.12}$$

Here, the trace on the RHS is understood in the Hilbert space l^2.

Note that l^2 is an infinite-dimensional space, we have to truncate the number of eigenfuntions to a finite number $N \in \mathbb{N}$. In this case, define the energy matrix $\boldsymbol{E}_N \in \mathbb{R}^{N\times N}$ by

$$\boldsymbol{E}_N = \mathrm{diag}\{E_0, E_1, \cdots, E_{N-1}\} \in \mathbb{R}^{N\times N}, \tag{D.13}$$

the potential matrix $\boldsymbol{V}_N^a \in \mathbb{R}^{N\times N}$ by

$$(\boldsymbol{V}_N^a)_{n,m} = \langle \psi_n | V^a | \psi_m \rangle\,, \quad n,m = 0,1,\cdots,N-1, \tag{D.14}$$

and the observable matrix $\boldsymbol{O}_N \in \mathbb{R}^{N\times N}$ by

$$(\boldsymbol{O}_N^a)_{n,m} = \langle \psi_n | V^a | \psi_m \rangle \quad n,m = 0,1,\cdots,N-1, \tag{D.15}$$

then the quantum thermal average $\langle O(\hat{q})\rangle_\beta$ is approximated by

$$\langle O(\hat{q})\rangle_\beta = \lim_{N\to\infty} \frac{\mathrm{Tr}[e^{-\beta(\boldsymbol{E}_N+\boldsymbol{V}_N^a)}\boldsymbol{O}]}{\mathrm{Tr}[e^{-\beta(\boldsymbol{E}_N+\boldsymbol{V}_N^a)}]}, \tag{D.16}$$

where the trace terms in the RHS of (D.16) can be computed directly.

Acknowledgement

The work of Z. Zhou is partially supported by the National Key R&D Program of China (Project No. 2020YFA0712000, 2021YFA1001200), and the National Natural Science Foundation of China (Grant No. 12031013, 12171013).

The authors would like to thank Haitao Wang (SJTU), Xin Chen (SJTU), Lei Li (SJTU), He Zhang (PKU), Lei Zhang (PKU) and Zexing Li (Cambridge) for the helpful discussions. The numerical experiments are supported by High-performance Computing Platform of Peking University.

References

[1] Kerson Huang. *Statistical mechanics*. John Wiley & Sons, 2008.

[2] Donald A McQuarrie. *Statistical mechanics*. Sterling Publishing Company, 2000.

[3] Peter Atkins, Peter William Atkins, and Julio de Paula. *Atkins' physical chemistry*. Oxford university press, 2014.

[4] Neil W Ashcroft and N David Mermin. *Solid state physics*. Cengage Learning, 2022.

[5] Subir Sachdev. Quantum phase transitions. *Physics world*, 12(4):33, 1999.

[6] Steven A Orszag. Comparison of pseudospectral and spectral approximation. *Studies in Applied Mathematics*, 51(3):253–259, 1972.

[7] Jie Shen, Tao Tang, and Li-Lian Wang. *Spectral methods: algorithms, analysis and applications*, volume 41. Springer Science & Business Media, 2011.

[8] William H Miller. Path integral representation of the reaction rate constant in quantum mechanical transition state theory. *The Journal of Chemical Physics*, 63(3):1166–1172, 1975.

[9] Ian R Craig and David E Manolopoulos. Chemical reaction rates from ring polymer molecular dynamics. *The Journal of chemical physics*, 122(8):084106, 2005.

[10] Xuecheng Tao, Philip Shushkov, and Thomas F Miller. Microcanonical rates from ring-polymer molecular dynamics: Direct-shooting, stationary-phase, and maximum-entropy approaches. *The Journal of Chemical Physics*, 152(12), 2020.

[11] Gregory A Voth. Feynman path integral formulation of quantum mechanical transition-state theory. *The Journal of Physical Chemistry*, 97(32):8365–8377, 1993.

[12] Edit Mátyus, David J Wales, and Stuart C Althorpe. Quantum tunneling splittings from path-integral molecular dynamics. *The Journal of chemical physics*, 144(11):114108, 2016.

[13] CL Vaillant, DJ Wales, and SC Althorpe. Tunneling splittings from path-integral molecular dynamics using a langevin thermostat. *The Journal of chemical physics*, 148(23):234102, 2018.

[14] Scott Habershon, David E Manolopoulos, Thomas E Markland, and Thomas F Miller III. Ring-polymer molecular dynamics: Quantum effects in chemical dynamics from classical trajectories in an extended phase space. *Annual review of physical chemistry*, 64:387–413, 2013.

[15] Ian R Craig and David E Manolopoulos. Quantum statistics and classical mechanics: Real time correlation functions from ring polymer molecular dynamics. *The Journal of chemical physics*, 121(8):3368–3373, 2004.

[16] Seogjoo Jang and Gregory A Voth. A derivation of centroid molecular dynamics and other approximate time evolution methods for path integral centroid variables. *The Journal of chemical physics*, 111(6):2371–2384, 1999.

[17] Jianshu Cao and Gregory A Voth. The formulation of quantum statistical mechanics based on the feynman path centroid density. iii. phase space formalism and analysis of centroid molecular dynamics. *The Journal of chemical physics*, 101(7):6157–6167, 1994.

[18] J Cao and GJ Martyna. Adiabatic path integral molecular dynamics methods. ii. algorithms. *The Journal of chemical physics*, 104(5):2028–2035, 1996.

[19] Mark E Tuckerman, Dominik Marx, Michael L Klein, and Michele Parrinello. Efficient and general algorithms for path integral car–parrinello molecular dynamics. *The Journal of chemical physics*, 104(14):5579–5588, 1996.

[20] Thomas E Markland and David E Manolopoulos. An efficient ring polymer contraction scheme for imaginary time path integral simulations. *The Journal of chemical physics*, 129(2):024105, 2008.

[21] Bruce J Berne and D Thirumalai. On the simulation of quantum systems: path integral methods. *Annual Review of Physical Chemistry*, 37(1):401–424, 1986.

[22] Jianfeng Lu, Yulong Lu, and Zhennan Zhou. Continuum limit and preconditioned langevin sampling of the path integral molecular dynamics. *Journal of Computational Physics*, 423:109788, 2020.

[23] Roman Korol, Jorge L Rosa-Raíces, Nawaf Bou-Rabee, and Thomas F Miller III. Dimension-free path-integral molecular dynamics without preconditioning. *The Journal of chemical physics*, 152(10):104102, 2020.

[24] James Glimm and Arthur Jaffe. *Quantum physics: a functional integral point of view*. Springer Science & Business Media, 2012.

[25] Jian Liu, Dezhang Li, and Xinzijian Liu. A simple and accurate algorithm for path integral molecular dynamics with the langevin thermostat. *The Journal of chemical physics*, 145(2):024103, 2016.

[26] Dominique Bakry, Ivan Gentil, Michel Ledoux, et al. *Analysis and geometry of Markov diffusion operators*, volume 103. Springer, 2014.

[27] Pierre Monmarché. Hypocoercivity in metastable settings and kinetic simulated annealing. *Probability Theory and Related Fields*, 172(3-4):1215–1248, 2018.

[28] Pierre Monmarché. Generalized Γ calculus and application to interacting particles on a graph. *Potential Analysis*, 50(3):439–466, 2019.

[29] Dominik Marx and Michele Parrinello. Ab initio path integral molecular dynamics: Basic ideas. *The Journal of chemical physics*, 104(11):4077–4082, 1996.

[30] Randall W Hall and Bruce J Berne. Nonergodicity in path integral molecular dynamics. *The Journal of chemical physics*, 81(8):3641–3643, 1984.

[31] Stephen D Bond, Brian B Laird, and Benedict J Leimkuhler. On the approximation of feynman–kac path integrals. *Journal of Computational Physics*, 185(2):472–483, 2003.

[32] Nawaf Bou-Rabee and Andreas Eberle. Two-scale coupling for preconditioned hamiltonian monte carlo in infinite dimensions. *Stochastics and Partial Differential Equations: Analysis and Computations*, 9:207–242, 2021.

[33] Roman Korol, Nawaf Bou-Rabee, and Thomas F Miller III. Cayley modification for strongly stable path-integral and ring-polymer molecular dynamics. *The Journal of chemical physics*, 151(12):124103, 2019.

[34] Timothy JH Hele, Michael J Willatt, Andrea Muolo, and Stuart C Althorpe. Boltzmann-conserving classical dynamics in quantum time-correlation functions:"matsubara dynamics". *The Journal of Chemical Physics*, 142(13):134103, 2015.

[35] Timothy JH Hele, Michael J Willatt, Andrea Muolo, and Stuart C Althorpe. Communication: Relation of centroid molecular dynamics and ring-polymer molecular dynamics to exact quantum dynamics. *The Journal of Chemical Physics*, 142(19):191101, 2015.

[36] Kenneth A Jung, Pablo E Videla, and Victor S Batista. Ring-polymer, centroid, and mean-field approximations to multi-time matsubara dynamics. *The Journal of Chemical Physics*, 153(12):124112, 2020.

[37] Yu Cao, Jianfeng Lu, and Lihan Wang. On explicit l^2-convergence rate estimate for underdamped langevin dynamics. *arXiv preprint arXiv:1908.04746*, 2019.

[38] Cédric Villani. Hypocoercivity. *arXiv preprint math/0609050*, 2006.

[39] Michael Reed. *Methods of modern mathematical physics: Functional analysis*. Elsevier, 2012.

[40] Ben Leimkuhler and Charles Matthews. Molecular dynamics. *Interdisciplinary applied mathematics*, 39:443, 2015.

[41] Steven A Orszag. Comparison of pseudospectral and spectral approximation. *Studies in Applied Mathematics*, 51(3):253–259, 1972.

[42] John C Light and Tucker Carrington. Discrete-variable representations and their utilization. *Advances in Chemical Physics*, 114:263–310, 2007.

[43] Peter D Lax. *Functional analysis*, volume 55. John Wiley & Sons, 2002.