

# Harmonic Oscillator Staging for Efficient Path Integral Simulations

Sabry G. Moustafa<sup>\*,†</sup> and Andrew J. Schultz<sup>‡</sup>

<sup>†</sup>*Department of Engineering Science, Trinity University, San Antonio, Texas 78212, USA*

<sup>‡</sup>*Department of Chemical and Biological Engineering, University at Buffalo, The State University of New York, Buffalo, New York 14260-4200, USA*

E-mail: smoustaf@trinity.edu

arXiv:2404.12551v1 [cond-mat.stat-mech] 19 Apr 2024

## Abstract

Imaginary-time path integral (PI) is a rigorous quantum mechanical tool to compute static properties at finite temperatures. However, the stiff nature of the internal PI modes poses a sampling challenge. This is commonly tackled using staging coordinates, in which the free particle (FP) contribution of the PI action is diagonalized. We introduce novel and simple staging coordinates that diagonalize the entire action of the harmonic oscillator (HO) model, rendering it efficiently applicable to systems with a harmonic character, such as quantum oscillators and crystals. Unlike FP staging, the HO staging provides a unique treatment of the centroid mode. We provide implementation schemes for PIMC and PIMD simulations in NVT ensemble. Sampling efficiency is assessed in terms of the precision and accuracy of estimating the energy and heat capacity of a one-dimensional HO, along with a harmonic+quartic anharmonic oscillator (AO) to investigate anharmonic effects. In PIMC, the HO coordinates propose collective moves that perfectly sample the HO contribution, then (for AO) the residual anharmonic term is sampled using standard Metropolis method. This results in a high acceptance rate and, hence, high precision, in comparison to the FP staging. In PIMD, the HO coordinates naturally prescribe definitions for the fictitious masses, yielding equal frequencies of all modes when applied to the HO model. This allows for a substantially larger time step sizes relative to standard staging, without affecting accuracy nor integrator stability. For completeness, we also present results using normal mode (NM) coordinates, based on both HO and FP models. While staging and NM coordinates show similar performance (for FP or HO), staging is computationally cheaper due to its linear, rather than quadratic, scaling with the number of beads. The simplicity and the enhanced sampling gained by the HO coordinates open avenues for efficient estimation of nuclear quantum effects in more complex systems, such as molecular bonds and quantum crystals.

## 1 INTRODUCTION

Nuclear quantum effects of distinguishable particles, such as zero-point energy and tunneling, play an important role when the thermal energy becomes smaller than the spacing between quantum energy levels. This is especially relevant to lightweight elements (e.g., hydrogen and helium) and/or high frequency (stiff) modes.<sup>1</sup> For crystalline systems, an often adopted quantitative threshold for these effects is the Debye temperature, which is proportional to the maximum phonon frequency present.<sup>2</sup> In these cases, quantum treatment for thermodynamic properties is crucial. According to quantum statistical mechanics, the canonical partition function of a system at a temperature  $T$  is given as the trace of its density matrix,  $Z = \text{tr} \left( e^{-\beta \hat{H}} \right)$ , where  $\hat{H}$  is the Hamiltonian operator and  $\beta$  is the reciprocal of the thermal energy,  $k_B T$ , with  $k_B$  the Boltzmann constant. However, except for simple cases, such as free particles and harmonic oscillators, it is challenging to determine  $Z$  analytically. In this regard, Feynman’s imaginary-time path integral (PI) formulation provides a numerically tractable alternative.<sup>3</sup> In this framework, the quantum system is mapped onto an extended classical isomorphism, made of  $n$  replicas (imaginary time slices) of the original system. These images are connected consecutively through harmonic springs, in a closed ring-polymer arrangement of  $n$  beads. The true partition function is then recovered in the  $n \rightarrow \infty$  limit.

The “standard” path integral primitive action uses the free-particle (FP) density matrix as a reference and adopts a second-order Trotter factorization. In this approximation, the partition function of a system of  $N$  distinguishable particles, occupying  $d$ -dimensional space, is given by the following integral over a  $dNn$  extended phase space,<sup>4</sup>

$$Z(\beta) = \left( \frac{mn}{2\pi\hbar^2\beta} \right)^{dNn/2} \int d\mathbf{x} \exp(-\beta V(\mathbf{x}, \beta)), \quad (1a)$$

$$V(\mathbf{x}, \beta) = \sum_{i=0}^{n-1} \frac{1}{2} \frac{m\omega_n^2}{n} (\mathbf{x}_i - \mathbf{x}_{i-1})^2 + \sum_{i=0}^{n-1} \frac{1}{n} U(\mathbf{x}_i), \quad (1b)$$

where  $m$  is the atom mass,  $\hbar \equiv h/2\pi$  is the reduced Planck constant,  $\omega_n \equiv n/\beta\hbar$ , and  $\mathbf{x}$  is a vector of all the  $dNn$  beads coordinates, with  $\mathbf{x}_i$  representing the coordinates of the  $i^{\text{th}}$  replica (of length  $dN$ ). Here,  $V$  represents a  $T$ -dependent effective potential, which comprises the intermolecular potentials  $U(\mathbf{x}_i)$ , from each replica, and the kinetic term, rep-

resented by FP harmonic interactions of identical spring constants,  $m\omega_n^2/n$ .

In dense phases,  $\beta\hbar\omega_{\text{max}}$  is often used to gauge “quantumness” level, where  $\omega_{\text{max}}$  is the maximum angular frequency in the system (Debye frequency in crystals).<sup>2</sup> On the other hand, numerical convergence in PI is achieved by setting the number of beads equal to multiples of the quantumness, such that  $n = C\beta\hbar\omega_{\text{max}}$ , with  $C$  being a constant greater than one.<sup>1</sup> Hence, based on the definition of  $\omega_n$ , the spring constant of the kinetic term,  $m\omega_n^2/n$ , is stiffer than the maximum intermolecular counterpart,  $m\omega_{\text{max}}^2/n$ , hence,  $\omega_n \gg \omega_{\text{max}}$  at convergence. The relatively stiff harmonic interactions are known to cause sampling challenges in both PIMD and PIMC simulations, using Cartesian coordinates.<sup>5</sup> For PIMD, ergodicity problems arise due to the dominant harmonic behavior, which has a poor energy exchange between different modes.<sup>6</sup> Nevertheless, this is usually alleviated using specific thermostating methods, such as Langevin and Nose-Hoover chain thermostats. More importantly, small time steps are necessary in order to resolve the fast stiff modes, which clearly introduces slow statistical convergence. For PIMC, on the other hand, the local nature of the MC moves (one bead at a time) and their small size (to handle stiff potentials) cause sampling inefficiency.<sup>5</sup> Since  $\omega_n$  is proportional to  $n$ , then these issues are more relevant at low temperatures, where more beads are needed for convergence.

These challenges are primarily due to using Cartesian coordinates, which are inherently local and, hence, do not account for the collective character of the system modes. A natural solution is then to transform the coordinates to a new set compatible with the PI action. The standard approaches are either to use staging<sup>5,7</sup> or normal mode (NM)<sup>8,9</sup> coordinates, which diagonalize the kinetic contribution of the PI action (i.e., FP model). In application to PIMD, fictitious (or, dynamic) masses are assigned to the transformed coordinates, in an attempt to narrow the time scale differences and, hence, improve sampling. The system dynamics will clearly change; however, this does not affect static properties, such as energy and heat capacity. For PIMC, a collective move of all beads associated with a given atom is proposed according to independent Gaussian distributions, which describe the kinetic term in the transformed coordinates. Then, the move is either accepted or rejected according to the usual Metropolis scheme, applied to the residual (off kinetic) energy contribution. Although

both coordinates diagonalize the same function, the appealing feature of staging is its  $\mathcal{O}(n)$  computational scaling, in contrast to  $\mathcal{O}(n^2)$  for the NM case.

A common limitation of the coordinates lies in the fact that both coordinates describe only the internal degrees of freedom within the ring-polymer. However, traditional approaches are typically adopted to describe the external degree of freedom, which is represented by the centroid and a bead position for the case of NM and staging coordinates, respectively. For example, a typical practice in PIMD simulation is to assign the physical mass to that mode. On the other hand, for PIMC, the external degree of freedom is sampled separately using a translation move of the whole ring-polymer, according to the standard Metropolis scheme. Clearly, both of these choices are not system-dependent. However, taking into account the nature of the system being considered could, in principle, improve sampling efficiency.

We consider applications to systems with harmonic character, such as quantum oscillators. For crystals, the atomic vibrations could be approximated as independent oscillators, forming an Einstein crystal. In these cases, the harmonic contribution could be described using the known action of the harmonic oscillator (HO) model. In this study, we introduce a set of staging and NM coordinates, which diagonalize the entire HO action rather than merely the kinetic term, as adopted in the FP methods. Hence, we denote these methods as ‘‘HO staging’’ and ‘‘HO NM’’ coordinates. Unlike FP approaches, the new coordinates represent both internal and external degrees of freedom and, hence, provide a better description of the real system. In application to PIMD, the approach provides a natural prescription for the fictitious masses, including those of the centroid mode. Consequently, all modes fluctuate on the same time scale in the case of HO model, or nearly similar time scales with anharmonic systems. This, then, would allow for using relatively large time step sizes, without affecting the accuracy nor the MD integrator stability. This, in turn, results in higher precision estimates, for a given number of steps. For PIMC, the HO coordinates prescribes a collective move that samples perfectly (100% acceptance) the HO contribution of the PI action. Unlike FP moves, the centroid degree of freedom is inherently included with the new approach. The move is then accepted/rejected using the standard Metropolis criterion, according to the residual, anharmonic, contribution. This would then result in a high overall acceptance rate and, hence, higher precision in measured properties.

To assess the performance, we apply the HO coordinates approach to measure the energy and heat capacity of a one-dimensional HO, in both PIMC and PIMD simulations, using the centroid virial estimator. Despite the model simplicity, it shares the same challenges encountered by PI simulations (stiff harmonic bonds, multiple time scales, nonergodicity, etc.). In addition, the model has an exact solution, which is used as a means to check accuracy. However, the effect of anharmonicity was also investigated using an anharmonic oscillator (AO), in which the HO is augmented by a quartic term. The statistical uncertainty in the measured PI estimators is used as a gauge to the sampling efficiency. In all cases, the HO coordinates show superior precision relative to existing FP coordinates. This is a manifestation of the method’s high acceptance rate and large time step size allowed for PIMC and PIMD simulations, respectively. Similar to the FP method, both HO staging and NM coordinates show similar performance. We also explore the dependence of sampling efficiency on the number of beads,  $n$ . A typi-

cal behavior of the centroid virial estimator is observed, in which the statistical uncertainty shows no sensitivity to the number of beads after some value (significantly smaller than the  $n = C\beta\hbar\omega_{\max}$  converged value).

It is worth to point out to another, less-common, second-order factorization, which uses the HO (rather than FP) propagator as a reference to formulate the discretized PI.<sup>10,11</sup> The primary advantage of this method is its relatively small (or non for HO) finite-size effects compared to primitive PI (Eq. 1). Nevertheless, since the ultimate interest is in the continuum limit ( $n \rightarrow \infty$ ), extrapolation is required in both PI formulations. The primitive PI formulation, however, offers a few advantages over the HO-based PI method, prompting our choice to it. For example, in addition to its simplicity, the HO spring constant is independent of temperature, unlike the case with the HO-based PI.<sup>11</sup> This adds complexity to the analysis, especially when deriving properties based on temperature derivatives, such as heat capacity (considered here).

The structure of this paper is as follows. In Sec. 2, we first introduce the HO staging and normal mode coordinates formulation as a approach to diagonalize the HO action. This is followed by providing schemes for implementing these coordinates in PIMC and PIMD simulations. For PIMD, we define the new fictitious masses associated with the HO coordinates. We then provide formulas for the total energy and heat capacity, based on the centroid virial estimator, along with uncertainty analysis for both quantities. We follow this by detailed model and simulation parameters considered. In Sec.3, we apply the new coordinates to both HO and AO models and show performance results based on both PIMC and PIMD techniques. We then conclude in Sec. 4 with a summary the framework and its outcomes, along with potential future directions.

## 2 FORMALISM AND METHODS

### 2.1 Harmonic Oscillator Staging

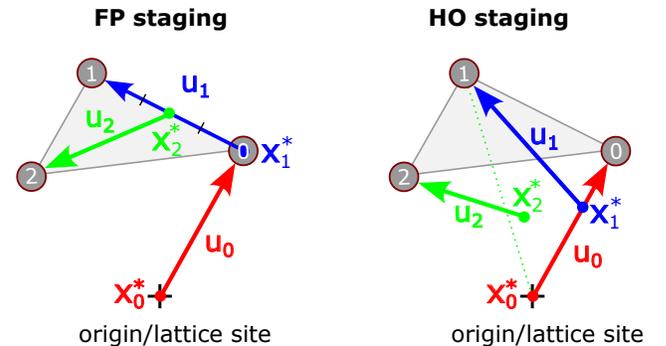


Figure 1: A simplified example of the FP (left) and HO (right) staging coordinates  $u_i$  (arrows) for a three-beads system. The instantaneous equilibrium position  $x_0^*$  is always located at the global minimum (origin/lattice position). For  $i > 0$ ,  $x_i^*$  for the FP staging lays on the line connecting beads 0 and  $i - 1$  according to Eq. 9, whereas, for HO staging it is contained within the triangle connecting the origin/lattice position and beads 0 and  $i - 1$ , according to by Eq. 5. For example,  $x_2^*$  is located within the origin-0-1 triangle, with the 0-1 side is connected by a dashed green line for clarity.

We will use the HO model to derive a new set of staging coordinates relevant to systems with a harmonic char-

acter, such as quantum oscillators and crystals. For simplicity of notation, we will develop the formulation for a one-dimensional HO. Extension to an arbitrary system size and/or dimensions is trivial due to the non-interacting nature of the oscillators, in all directions. According to Eq. 1b, the effective potential of the HO model is given by,

$$V(\mathbf{x}, \beta) = \sum_{i=0}^{n-1} \frac{1}{2} \frac{m\omega_n^2}{n} (x_i - x_{i-1})^2 + \sum_{i=0}^{n-1} \frac{1}{2} \frac{m\omega^2}{n} x_i^2, \quad (2)$$

where  $x_i$  is the displacement of the  $i^{\text{th}}$  bead from its equilibrium position (lattice site for crystals), and  $\omega$  is the HO angular frequency. The objective now is to diagonalize this coupled form within the staging framework. To accomplish this, we define the following set of generalized staging coordinates:

$$u_i(x_0, x_1, \dots, x_i) \equiv x_i - x_i^*(x_0, x_1, \dots, x_{i-1}), \quad i > 0 \quad (3)$$

where  $x_i^*$  is an instantaneous ‘‘equilibrium’’ position, solely dependent on the positions of the previous beads ( $0, \dots, i-1$ ). We define the zero staging coordinate as  $u_0 = x_0$  (hence,  $x_0^* = 0$ ), which acts as an external mode, somewhat similar to the centroid. We then use these coordinates to express the effective PI potential (Eq. 2) in the following decoupled form of independent harmonic oscillators:

$$V(\mathbf{u}, \beta) = \sum_{i=0}^{n-1} \frac{1}{2} k_i u_i^2, \quad (4)$$

where  $k_i$  is the staging spring constant associated with stage  $i$ . This form has two unknowns, for each mode,  $x_i^*$  and  $k_i$ , which we determine in the Appendices. The  $x_i^*$  positions are determined through first derivative of  $V$  with respect to coordinates (Appendix A),

$$x_i^*(x_0, x_{i-1}) = \frac{\sinh(\alpha)x_0 + \sinh((n-i)\alpha)x_{i-1}}{\sinh((n-i+1)\alpha)}, \quad i > 0 \quad (5)$$

while the spring constants are obtained through second derivatives (Appendix B),

$$k_i = \frac{m\omega_n^2}{n} \begin{cases} 2 \sinh(\alpha) \tanh\left(\frac{n\alpha}{2}\right), & i = 0 \\ \frac{\sinh((n-i+1)\alpha)}{\sinh((n-i)\alpha)}, & i > 0 \end{cases} \quad (6)$$

In Appendix C, we provide a verification on the equivalence between the original potential (Eq. 2) and the new staging diagonalization (Eq. 4).

According to Eqs. 3 and 5, the  $\mathbf{x} \rightarrow \mathbf{u}$  transformation is given by

$$u_i = x_i - \frac{\sinh(\alpha)x_0 + \sinh((n-i)\alpha)x_{i-1}}{\sinh((n-i+1)\alpha)}, \quad i > 0 \quad (7)$$

where, again,  $u_0 = x_0$ . Then, the corresponding inverse ( $\mathbf{x} \leftarrow \mathbf{u}$ ) transformation is

$$x_i = u_i + \frac{\sinh(\alpha)u_0 + \sinh((n-i)\alpha)x_{i-1}}{\sinh((n-i+1)\alpha)}, \quad i > 0 \quad (8)$$

where  $x_n = x_0$ , and we used that fact that  $x_0 = u_0$ . Note that, while the forward transformation is non-recursive, the inverse one is recursive (starting from  $i = 1$  to  $n-1$ ). Both transformations are simple operations of element-by-element multiplications, hence the CPU cost scales with the number of beads as  $\mathcal{O}(n)$ . On the other hand, as discussed below,

the NM transformations are given as matrix-vector multiplications, which yield a higher computational cost of  $\mathcal{O}(n^2)$ .

It is interesting to note that the functional form of the HO staging is identical to that of the harmonic staging associated with the PI action based on the HO reference.<sup>11</sup> The only difference is the  $\alpha$  form, which reduces to our formula in the limit of large  $n$ . The similarity is attributed to the fact that both approaches diagonalize the entire HO action, with only different spring constants.

Contrary to the HO staging, the staging as known in literature (denoted here as ‘‘FP staging’’) diagonalizes only the free particle (kinetic) contribution of the PI action (first term of Eq. 1b). Interestingly, the FP staging is a special case of the HO staging when taking the  $\omega \rightarrow 0$  (or,  $\alpha \rightarrow 0$ ) limit of Eqs. 5 and 6,

$$u_i \equiv x_i - x_i^* = x_i - \frac{x_0 + (n-i)x_{i-1}}{n-i+1}, \quad i > 0 \quad (9)$$

where  $u_0 = x_0$ . From this, the inverse ( $\mathbf{x} \leftarrow \mathbf{u}$ ) transformation is given by

$$x_i = u_i + \frac{u_0 + (n-i)x_{i-1}}{n-i+1}, \quad i > 0 \quad (10)$$

where  $x_n = x_0$ . Similarly, the FP staging spring constants are given in the limit of Eq. 6 as

$$k_i = \frac{m\omega_n^2}{n} \frac{n-i+1}{n-i}, \quad i > 0, \quad (11)$$

where  $k_0 = 0$ .

Note that we adopt a counterclockwise approach ( $0, 1, \dots, n-1$ ) to derive the HO and FP staging coordinates (see Fig. 8 in Appendix A). However, it is trivial to recover the clockwise version by simply replacing the  $x_{i-1}$  position by  $x_{i+1}$ , and  $n-i$  by  $i$  (not including subscribes of the positions). In Appendix D we provide the clockwise version of both HO and FP stagings, both coordinates and spring constants. We emphasize that both versions are equivalent since they both diagonalize the same HO potential.

Figure 1 provides an illustrative example of a three-beads system, using both FP and HO staging methods. For the FP staging case, according to Eq. 9,  $x_i^*$  corresponds to an interpolation along the line connecting beads 0 and  $i-1$ . On the other hand, for the HO staging method, the  $x_i^*$  location is within the triangle connecting the origin and beads 0 and bead  $i-1$ , as can be recognized from Eq. 5.

## 2.2 Harmonic Oscillator Normal Modes

The HO effective potential (Eq. 2) can be expressed in terms of normal mode (NM) coordinates,  $q_k$ , in the following uncoupled form<sup>12</sup>

$$V(\mathbf{q}, \beta) = \sum_{k=0}^{n-1} \frac{1}{2} \Lambda_k q_k^2, \quad (12a)$$

$$\Lambda_k \equiv n\lambda_k = m\omega^2 + 4m\omega_n^2 \sin^2\left(\frac{\pi k}{n}\right), \quad k = 0, 1, \dots, n-1 \quad (12b)$$

where  $\lambda_k$  are the eigenvalues of the following  $n \times n$  Hessian matrix  $H$  of second derivatives:

$$H = \frac{m\omega_n^2}{n} \begin{pmatrix} c & -1 & & -1 \\ -1 & c & -1 & \\ & & \ddots & \\ -1 & & -1 & c & -1 \\ & & & -1 & c \end{pmatrix} \quad (13)$$

where  $c \equiv 2 + \frac{\omega^2}{\omega_n^2}$ , such that  $V = \frac{1}{2}\mathbf{x} \cdot H \cdot \mathbf{x}$ . Note that  $H$  is a symmetric circulant tridiagonal matrix.

The normal mode coordinates vector  $\mathbf{q}$  is related to the Cartesian coordinates vector  $\mathbf{x}$  (both of length  $n$ ) through,<sup>13</sup>

$$\mathbf{x} = \sqrt{n}A\mathbf{q} \quad (14)$$

where  $A$  is a  $n \times n$  matrix of the orthonormal eigenvectors associated with the Hessian matrix, hence the inverse is equal to the transpose,  $A^{-1} = A^t$ . Note that the zero-frequency mode is just the centroid coordinate,  $q_0 = x_c$ . Adopting the real-valued representation of the normal modes, the  $jk$  component of  $A$  is given by

$$A_{jk} = \frac{1}{\sqrt{n}} \begin{cases} 1, & k = 0 \\ \sqrt{2} \sin\left(\frac{2\pi jk}{n}\right), & -n_2 \leq k < 0 \\ \sqrt{2} \cos\left(\frac{2\pi jk}{n}\right), & 0 < k \leq n_2 \\ (-1)^j, & k = n_2 + 1 \text{ (even } n) \end{cases} \quad (15)$$

where  $n_2 = (n-1)/2$  for odd  $n$ , and  $n_2 = (n-2)/2$  for even  $n$ .

A special case of the HO NM coordinates is the FP NM method, wherein only the kinetic contribution of Eq. 2 is considered in the Hessian matrix; i.e.,  $c = 2$  in Eq. 13. Accordingly, the FP NM eigenvalues (multiplied by  $n$ ) are obtained by setting  $\omega = 0$  in Eq. 12b,

$$\Lambda_k = 4m\omega_n^2 \sin^2\left(\frac{\pi k}{n}\right), k = 0 \dots n-1. \text{ (FP NM)} \quad (16)$$

However, the associated eigenvectors are the *same* as the HO NM case, due to the preservation of the structure of the Hessian matrix.

Note that, expressions for  $x_i^*$  (Eq. 5),  $k_i$  (Eq. 6), and  $\Lambda_k$  (Eq. 12b) are universal, regardless of the number of dimensions  $d$  and/or atoms  $N$ . Hence, in the analysis below, we apply the HO staging and NM methods to sample coordinates according to the PIMC and PIMD approaches, in a general space.

### 2.3 Application to PIMC

In the PIMC simulations, we sample configurations in the NVT canonical ensemble. We first introduce the sampling scheme for the HO coordinates, then present the traditional procedure for the FP coordinates case as known in literature. The HO staging and NM coordinates are sampled according to the following splitting of the total probability distribution,

$$e^{-\beta V} = e^{-\beta V^{\text{HO}}} \times e^{-\beta V^{\text{AH}}}, \quad (17)$$

where  $V^{\text{HO}}$  is the harmonic effective potential (Eq. 2) and  $V^{\text{AH}}$  is the anharmonic contribution,

$$V^{\text{AH}} \equiv V - V^{\text{HO}} = \frac{1}{n} \sum_{i=0}^{n-1} \left( U(\mathbf{x}_i) - \frac{1}{2}m\omega^2 \mathbf{x}_i^2 \right). \quad (18)$$

where  $\mathbf{x}_i$  is a vector (of length  $dN$ ) containing the beads positions associated with the  $i^{\text{th}}$  replica. According to this splitting, the sampling consists of two steps: a collective move proposal of the  $n$  beads according to the HO contribution (Eq. 4 for staging, or Eq. 12 for NM), followed by accept/reject the move according to the anharmonic contribution (Eq. 18). The following algorithm describes this procedure for the HO coordinates:

1. For crystals of  $N$  atoms, pick a random atom, or skip to step “2” for quantum oscillators.
2. For each dimension, draw  $n$  transformed coordinates ( $\mathbf{u}$  for staging, or  $\mathbf{q}$  for NM) according to a set of Gaussian distributions with, zero means and standard deviations of
$$\sigma_i = \frac{1}{\sqrt{\beta f_i}}, i = 0, \dots, n-1$$
where  $f_i = k_i$  for HO staging (Eq. 6) and  $f_i = \Lambda_i$  for the HO NM (Eq. 12b).
3. Transform back to the Cartesian coordinates, according to Eq. 8 for HO staging, or Eq. 14 for HO NM methods.
4. Use the standard Metropolis MC method to accept/reject the proposed move according to the *change* in the anharmonic potential (Eq. 18) between the old and new configurations,  $\Delta V^{\text{AH}}$ .
5. If the move is accepted, update the beads positions, otherwise use the old configuration.
6. Repeat.

For the case of crystals, the displacements from the lattice sites are used as the Cartesian coordinates.

This algorithm is applicable to the FP coordinates as well, with the following exceptions. First, in step 2, the non-zero collective modes are still sampled according to Eq. 4 for staging and Eq. 12 for NM, but using the FP force constants; i.e., Eq. 11 for staging and Eq. 16 for NM. This move results in a change in the center of mass (centroid), which we bring back to where it was after the move. Second, the centroid (zero-mode) is then sampled separately using a rigid translation move of the ring polymer (with a target acceptance of 50%). Both of these moves are proposed with an equal probability. Third, the inverse transformation (step 3) for the FP staging is done using Eq. 10. But same transformation is used for both HO and FP coordinates (Eq. 14). Lastly, the residual energy used in step 4 is the difference from the kinetic term

$$V^{\text{residual}} \equiv V - V^{\text{FP}} = \frac{1}{n} \sum_{i=0}^{n-1} U(\mathbf{x}_i), \quad (19)$$

which is just the external energy contribution.

### 2.4 Application to PIMD

Sampling configurations in PIMD is governed by fictitious dynamics, according to the following effective Hamiltonian in a general (staging or NM) phase space of coordinates  $q$  and momenta  $p$ :

$$H(q, p) = \sum_l \frac{p_l^2}{2m_l} + V(q, \beta), \quad (20)$$

where  $m_l$  is a fictitious mass (defined below) associated with the  $l$  degree of freedom and  $V$  is the effective PI potential (Eq. 4 for staging and Eq. 12a for NM). We run PIMD simulations in NVT canonical ensemble using Langevin thermostat, which we detail next.

#### 2.4.1 White noise Langevin thermostat

The Langevin dynamics equations of motion, in a general coordinate system, can be split into three parts,<sup>14,15</sup>

$$\begin{pmatrix} \dot{q}_l \\ \dot{p}_l \end{pmatrix} = \underbrace{\begin{pmatrix} p_l \\ m_l \\ 0 \end{pmatrix}}_A + \underbrace{\begin{pmatrix} 0 \\ F_l \end{pmatrix}}_B + \underbrace{\begin{pmatrix} 0 \\ -\gamma p_l + \sqrt{2m_l k_B T \gamma} \xi_l \end{pmatrix}}_O \quad (21)$$

where, for a given  $l$  degree of freedom,  $F_l \equiv -\partial V / \partial q_l$  is the effective force,  $\xi_l$  describes the Gaussian white noise, with  $\langle \xi_l(t) \rangle = 0$  and  $\langle \xi_l(t) \xi_{l'}(t') \rangle = \delta_{ll'} \delta(t' - t)$ , where  $\delta_{ll'}$  the Kronecker delta function, and  $\gamma$  is the thermostat friction coefficient (set to  $\gamma = \omega$  for optimum performance<sup>16</sup>). The ‘‘A’’ part is a free flight drift, while ‘‘O’’ (Ornstein-Uhlenbeck process) represents random kicks in momenta required to guarantee a canonical ensemble, both are solved exactly. On the other hand, the ‘‘B’’ contribution represents deterministic kicks in momenta, caused by the real system forces  $F_l$ , and it sets an upper bound on the Verlet time step size.<sup>14</sup>

Due to its efficiency, we adopt the BAOAB splitting scheme, in which a half time step is used to propagate the positions and momenta,<sup>14,16</sup>

$$B : p_l \left( t + \frac{\Delta t}{2} \right) = p_l(t) + F_l(t) \frac{\Delta t}{2} \quad (22a)$$

$$A : q_l \left( t + \frac{\Delta t}{2} \right) = q_l(t) + \frac{p_l \left( t + \frac{\Delta t}{2} \right) \Delta t}{m_l} \quad (22b)$$

$$O : p_l \left( t + \frac{\Delta t}{2} \right) \leftarrow p_l \left( t + \frac{\Delta t}{2} \right) \exp(-\gamma \Delta t) \quad (22c)$$

$$+ \sqrt{m_l k_B T (1 - \exp(-2\gamma \Delta t))} R_l(t) \quad (22d)$$

$$A : q_l(t + \Delta t) = q_l \left( t + \frac{\Delta t}{2} \right) + \frac{p_l \left( t + \frac{\Delta t}{2} \right) \Delta t}{m_l} \quad (22e)$$

$$B : p_l(t + \Delta t) = p_l \left( t + \frac{\Delta t}{2} \right) + F_l(t + \Delta t) \frac{\Delta t}{2} \quad (22f)$$

where  $R_l$  is a normal distribution random number, with  $\langle R_l \rangle = 0$  and  $\langle R_l R_{l'} \rangle = \delta_{ll'}$ .

We apply this algorithm to PIMD, in both staging and NM coordinates (HO and FP methods). In this application, the system evolves in the transformed coordinates according to the BAOAB procedure above. However, since the actual forces are defined in the Cartesian coordinates, we first perform inverse transformation of the coordinates ( $\mathbf{q} \rightarrow \mathbf{x}$ ), followed by force evaluation in the Cartesian space. These forces are then transformed to the staging/NM coordinate system, which is discussed in the next two sections for the staging and NM coordinates.

#### 2.4.2 Cartesian to staging force transformation

For a given atom, the force acting on bead  $j$  in the HO or FP staging coordinates is given by,

$$F_j = -\frac{\partial V}{\partial u_j}, \quad j = 0, 1, \dots, n-1 \quad (23)$$

Using the chain rule, these forces can be represented in terms of the Cartesian forces,  $F_j^{xyz} = -\partial V / \partial x_i$ , as

$$F_j = \sum_{i=j}^{n-1} \frac{\partial x_i}{\partial u_j} F_i^{xyz}, \quad (24)$$

where  $F_i^{xyz} = -\frac{m\omega_n^2}{n} (2x_i - x_{i-1} - x_{i+1}) - \frac{1}{n} \frac{\partial U(x_i)}{\partial x_i}$ , according to Eq. 1b. To evaluate Eq. 24, we need a closed form of the coordinate transformation, which can be obtained recursively using Eq. 8,

$$x_i(u_0, \dots, u_i) = \frac{\cosh\left(\left(\frac{n}{2} - i\right)\alpha\right)}{\cosh\left(\frac{n}{2}\alpha\right)} u_0 + \sum_{j=1}^i \frac{\sinh\left((n-i)\alpha\right)}{\sinh\left((n-j)\alpha\right)} u_j \quad (25)$$

Using this in Eq. 24, yields the following backward recursion relations for forces,

$$F_j = \begin{cases} \sum_{i=0}^{n-1} \frac{\cosh\left(\left(\frac{n}{2} - i\right)\alpha\right)}{\cosh\left(\frac{n}{2}\alpha\right)} F_i^{xyz}, & j = 0 \\ F_j^{xyz} + \frac{\sinh\left((n-j-1)\alpha\right)}{\sinh\left((n-j)\alpha\right)} F_{j+1}, & j = n-1, \dots, 1 \end{cases} \quad \text{(HO staging)} \quad (26)$$

with  $F_n = F_0$ .

On the other hand, the FP staging forces are given as a special case of the HO staging  $\omega \rightarrow 0$  (or,  $\alpha \rightarrow 0$ ) limit in Eq. 26,

$$F_j = \begin{cases} \sum_{i=0}^{n-1} F_i^{xyz}, & j = 0 \\ F_j^{xyz} + \frac{n-j-1}{n-j} F_{j+1}, & j = n-1, \dots, 1 \end{cases} \quad \text{(FP staging)} \quad (27)$$

Note that, this transformation is already known in literature,<sup>5,16</sup> but using the clockwise staging (see Appendix D).

#### 2.4.3 Cartesian to NM force transformation

Given that the eigenvectors for the NM coordinates using both the HO and FP methods are the same, it follows that the corresponding force transformations are also identical. Due to the nature of the NM transformations (Sec. 2.2), we adopt a matrix notation throughout. Similar to the staging case, we will focus on the beads coordinates associated with a given atom. The forces acting on the  $n$  beads, in the NM coordinates  $\mathbf{q}$ , are given by the following force vector,

$$\mathbf{F} = -\frac{\partial V}{\partial \mathbf{q}} \quad (28)$$

where both  $\mathbf{q}$  and  $\mathbf{F}$  vectors are of length  $dn$ . Using the chain rule, along with the  $\mathbf{q} \rightarrow \mathbf{x}$  transformation (Eq. 14), these forces are expressed in terms of the Cartesian coordinates forces  $\mathbf{F}^{xyz}$  as follows:

$$\mathbf{F} = \frac{\partial \mathbf{x}}{\partial \mathbf{q}} \cdot \mathbf{F}^{xyz} = \sqrt{n} A^t \mathbf{F}^{xyz} \quad (29)$$

where  $\mathbf{F}^{xyz} \equiv -\frac{\partial V}{\partial \mathbf{x}}$ .

It is worthwhile to emphasize that, despite the simplicity of the NM transformations, the computational cost of these force transformations is of order  $\mathcal{O}(n^2)$ , as mentioned earlier, due to the matrix-vector multiplication operations. This differentiates it from the staging method, which is only of order  $\mathcal{O}(n)$ , because of the recursive nature of the transformations (e.g., Eq. 26).

#### 2.4.4 Fictitious masses

Efficient PIMD simulations narrow the gap between the diverse time scales of PI systems by introducing a set of mode-based fictitious (or dynamic) masses.<sup>17</sup> This freedom of choosing the masses allows for using larger MD time steps, compared to the small time steps needed for the stiff internal PI vibrations. We define these masses in the HO staging and NM coordinates such that all modes (including the centroid) fluctuate on the *same* time scale for HO systems. However, for a meaningful comparison across different coordinates at a given time step  $\Delta t$ , we require this time scale to match the motion of the centroid in the Cartesian space. We set the physical mass of each bead to  $m/n$ ; hence, the centroid has a mass of  $m$  and oscillates at a frequency of  $\omega$ . To achieve this with the HO methods, we define the fictitious masses as  $m_i \equiv k_i/\omega^2$  for the HO staging (Eq. 6),

$$m_i = \frac{m\omega_n^2}{n\omega^2} \begin{cases} 2 \sinh(\alpha) \tanh\left(\frac{n\alpha}{2}\right), & i = 0 \\ \frac{\sinh((n-i+1)\alpha)}{\sinh((n-i)\alpha)}, & i = 1, 2, \dots, n-1 \end{cases} \quad (\text{HO staging}) \quad (30)$$

and as  $m_k \equiv \Lambda_k/\omega^2$  for the HO NM (Eq. 12b),

$$m_k = m \left[ 1 + 4 \frac{\omega_n^2}{\omega^2} \sin^2\left(\frac{\pi k}{n}\right) \right], \quad k = 0, 1, \dots, n-1 \quad (\text{HO NM}) \quad (31)$$

Since these choices unify the time scales for the HO model, anharmonicity would disturb this unity and could affect the allowed time step size, hence the sampling performance. However, still the performance is expected to be efficient relative to the FP methods, as supported by current results. These expectations are further discussed in the Results sections in for the AO model.

On the other hand, for the FP coordinates, we adopt the fictitious masses definitions as already given in the literature.<sup>7,16,18</sup> For FP staging and NM coordinates, these masses are given, respectively, by

$$m_i = m \begin{cases} 1, & i = 0 \\ \frac{n-i+1}{n-i}, & i > 0 \end{cases} \quad (\text{FP staging}) \quad (32)$$

and

$$m_k = m \begin{cases} 1, & k = 0 \\ 4n \sin^2\left(\frac{\pi k}{n}\right), & k > 0 \end{cases} \quad (\text{FP NM}) \quad (33)$$

Note that, for the FP NM case, the zero-mode represents the centroid coordinate; i.e.,  $q_0 = x_c$ , see Sec. 2.2. Therefore, the  $m_0 = m$  definition already yields a centroid oscillation of a frequency  $\omega$  in the real space, which satisfies the aforementioned criterion of matching the centroid motion. Unlike the NM case, all the staging  $u_i$  modes contribute to the centroid, which results in a complex dynamics of its motion. However, since the zero staging mode,  $u_0$ , defines the external Cartesian (absolute) coordinate of one bead, it is expected to somewhat resemble the centroid motion. In fact, we find this to be the case for temperatures above some intermediate value ( $\approx 0.4$ ), using PIMD simulations in NVE ensemble (not shown). This is to be expected as the internal interactions become stiffer with increasing the temperature (large  $n$ ), such that their contribution to the centroid becomes smaller in comparison to  $u_0$ . Thus, the  $m_0 = m$  choice with FP staging also yields (yet approximately) a centroid oscillation with a frequency  $\omega$ . However, at low temperatures, the centroid oscillates faster with complex dynamics, such that comparing performance with other methods, at a given  $\Delta t$ , not meaningful. However, in prac-

tice, interest is in determining the maximum allowable time step size that preserves accuracy.

## 2.5 Centroid Virial Estimator

We consider the total energy and heat capacity at constant volume as example properties to assess the sampling efficiency of both PIMC and PIMD simulations. We adopt the centroid virial estimator for both quantities, due to its known computational advantages over others (e.g., thermodynamic or virial estimators).<sup>19</sup> For a  $d$ -dimensional system of  $N$  species and  $n$  beads, the average total energy  $E$  and heat capacity  $C_V$  are given, respectively, as

$$E = \frac{dN}{2\beta} + \left\langle \sum_{i=0}^{n-1} \frac{1}{n} U(\mathbf{x}_i) - \frac{1}{2} \sum_{i=0}^{n-1} \mathbf{F}_i^0 \cdot (\mathbf{x}_i - \mathbf{x}_c) \right\rangle, \quad (34a)$$

$$\frac{C_V}{k_B\beta^2} = \frac{dN}{2\beta^2} + \text{Var}(E) + \frac{1}{4\beta} \left\langle 3 \sum_{i=0}^{n-1} \mathbf{F}_i^0 \cdot (\mathbf{x}_i - \mathbf{x}_c) - \sum_{i=0}^{n-1} (\mathbf{x}_i - \mathbf{x}_c) \cdot H_{ii}^0 \cdot (\mathbf{x}_i - \mathbf{x}_c) \right\rangle, \quad (34b)$$

where  $\mathbf{x}_c \equiv \frac{1}{n} \sum_{i=0}^{n-1} \mathbf{x}_i$  is a vector of the centroid coordinates associated with each atom (of length  $dN$ ). Moreover, the intermolecular forces (of length  $dN$ ) and Hessian matrix ( $dN \times dN$ ) associated with the  $i^{\text{th}}$  replica are given in terms of the intermolecular potential  $U$  as  $\mathbf{F}_i^0 = -\frac{1}{n} \frac{\partial U(\mathbf{x}_i)}{\partial \mathbf{x}_i}$  and  $H_{ii}^0 = \frac{1}{n} \frac{\partial^2 U(\mathbf{x}_i)}{\partial \mathbf{x}_i \partial \mathbf{x}_i}$ , respectively. The superscript “0” is used to denote the absence of internal harmonic interactions (i.e., kinetic term in Eq. 1b). Finally,  $\text{Var}(E) \equiv \langle E^2 \rangle - \langle E \rangle^2$  is the variance of the instantaneous energy (quantity being averaged in Eq. 34a).

## 2.6 Uncertainty Analysis

Uncertainties in the energy (Eq. 34a) and heat capacity (Eq. 34b) are estimated from single runs, using the block averaging technique.<sup>20</sup> This is straightforward to implement with the energy, since it is given as an average. However, the heat capacity is given as a variance and, hence, propagation of uncertainty is needed to express its uncertainty in terms of measured averages. According to Eq. 34b, the heat capacity can be written in this compact form:

$$C_V = \bar{Y} - \bar{E}^2, \quad (35)$$

where  $\bar{E}^2$  is the square of the block averaged energy, and  $\bar{Y}$  is a block average of the remaining terms. Propagating the uncertainty to the second-order in fluctuations, yields

$$\sigma^2(C_V) = \sigma^2(\bar{Y}) + 4\bar{E}^2 \sigma^2(\bar{E}) - 4\bar{E} \text{cov}(\bar{Y}, \bar{E}) \quad (36)$$

where  $\sigma^2(Y)$  and  $\text{cov}(Y, Z)$  are the variance and covariance, respectively. For a sufficiently large block size, the blocks can be considered independent, and the uncertainty can be written in terms of the statistics of the raw data,

$$\sigma^2(C_V) = \frac{1}{N_{\text{blocks}}} \left[ \sigma^2(Y) + 4\bar{E}^2 \sigma^2(E) - 4\bar{E} \text{cov}(Y, E) \right] \quad (37)$$

where  $N_{\text{blocks}}$  is the number of blocks (100 in our case). We verified the correctness of this formula against direct measurement of the standard deviation from 100 independent

runs.

## 2.7 Simulation Details

Two, one-dimensional, quantum oscillators are considered as prototypes: a harmonic oscillator (HO) and a harmonic+quartic anharmonic oscillator (AO), with the following potentials,

$$U^{\text{HO}} = \frac{1}{2}m\omega^2 x^2, \quad (38a)$$

$$U^{\text{AO}} = \frac{1}{2}m\omega^2 x^2 + k_4 x^4, \quad (38b)$$

where we set  $m = \omega = 1$  and  $k_4 = 0.1$  throughout this work. Despite their simplicity, these models have all the problems associated with PI simulations (e.g., nonergodicity, multiple time scales, CPU scaling with  $n$ , etc.). In addition, the HO has an exact energy form in the primitive PI action of  $n$  beads,<sup>9,21,22</sup>

$$E = \frac{\hbar\omega}{2} \frac{\coth\left(\frac{n\alpha}{2}\right)}{\sqrt{1 + \frac{1}{4}\epsilon^2}} \quad (39)$$

where  $\alpha$  and  $\epsilon$  are as defined above. The corresponding energy in the continuum limit ( $n \rightarrow \infty$ , or  $\epsilon \rightarrow 0$ ) is then given by,

$$E_\infty = \frac{\hbar\omega}{2} \coth\left(\frac{\beta\hbar\omega}{2}\right). \text{(continuum limit)} \quad (40)$$

The heat capacity at constant volume,  $C_V$ , can then be obtained from the first derivative of the energy with respect to temperature.

Both PIMC and PIMD simulations ran in NVT canonical ensemble, for  $10^7$  steps, after  $10^6$  steps of equilibration. A temperature range of  $T = 0.05 - 1.0$ , with an increment of 0.05, was considered, which shows significant quantum effects (see Fig. 2). For the PIMD simulations, the time step size was chosen for each method to be the largest possible value that did not sacrifice accuracy. Specific values for each case are given in the Results section. White noise Langevin (stochastic) thermostat was adopted with PIMD to ensure a canonical distribution, using the BAOAB integrator method.<sup>14,16</sup> The thermostat friction coefficient was set to  $\gamma = \omega$ , which is the common frequency of all modes in the HO staging and NM coordinates. Although this is the optimum value in the absence of external field,<sup>16</sup> current (and previous<sup>23</sup>) results did not show sensitivity to  $\gamma$  in the vicinity of this value.

Based on the convergence analysis with respect to the number of beads, we find that  $n = 20\beta\hbar\omega$  to yield statistically converged energy and heat capacity, for both the HO and AO models (see Appendix E). This corresponds to  $n = 20$  ( $T = 1.0$ ) and 400 ( $T = 0.05$ ). It is worthwhile to emphasize that the scaling factor ( $C = 20$ ) is only suitable for the level of precision attained here — larger  $n$  values would be needed for higher precision levels.

Both PIMC and PIMD simulations were performed using the Etomica simulation software,<sup>24</sup> which can be accessed at [https://github.com/etomica/etomica/tree/path\\_integral](https://github.com/etomica/etomica/tree/path_integral)

## 3 RESULTS AND DISCUSSION

To gain an insight on the degree of quantumness within the given temperature range ( $T = 0.05 - 1.0$ ), we present in

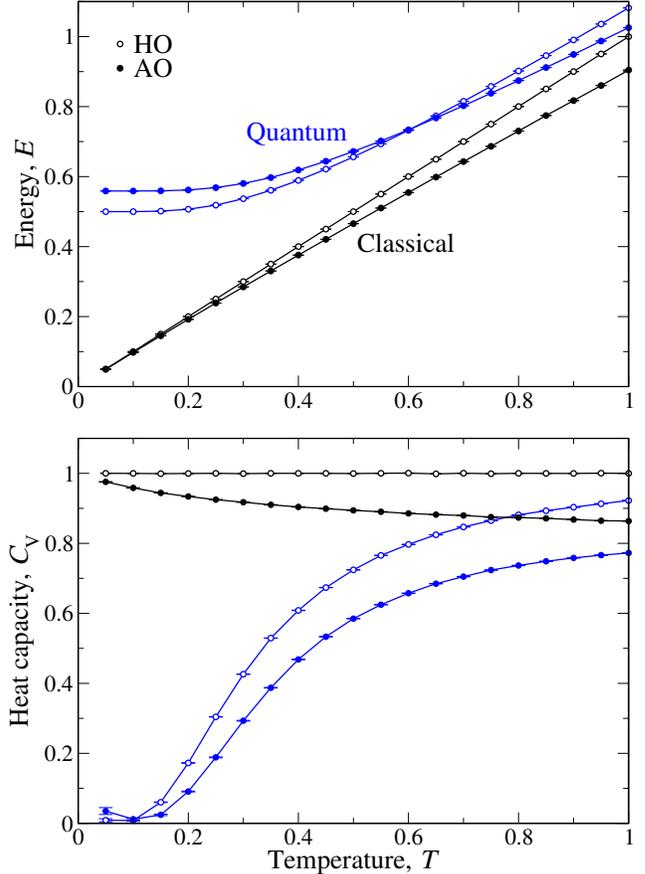


Figure 2: Temperature dependence of the energy  $E$  (top) and heat capacity at constant volume  $C_V$  (bottom) of a one-dimensional HO and AO models (Eqs. 38). The classical treatment ( $n = 1$ ) is presented to show the level of quantumness. Most error bars are smaller than the symbol size and lines join the data points as a guide to the eye. The data were generated from PIMC simulations using HO staging coordinates.

Fig. 2 the quantum and classical ( $n = 1$ ) representations of the total energy and heat capacity, for both the HO and AO models. The data were generated using PIMC simulation, in HO staging coordinates. We notice that the difference between the quantum and classical representations is substantial across the whole temperature range considered, with the maximum value at zero temperature (i.e., zero-point energy). We now use these highly quantum models to investigate the performance of the staging and NM coordinates, using both PIMC and PIMD simulation methods.

### 3.1 PIMC

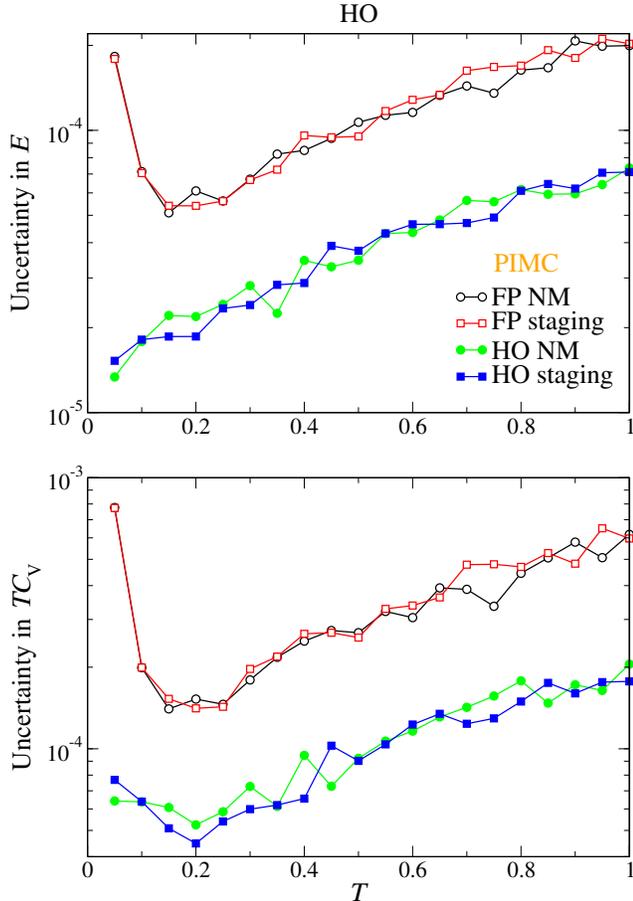


Figure 3: Temperature dependence of the PIMC statistical uncertainty in estimating the energy (top) and heat capacity multiplied by the temperature (bottom) of the HO model, using both HO and FP coordinates. Lines join the data points as a guide to the eye.

Figure 3 presents the PIMC statistical uncertainty in estimating the energy (top) and heat capacity multiplied by  $T$  (bottom) of the HO model, using both HO and FP methods. Results from the AO model (not shown) are qualitatively similar. The first observation to report is the statistically indistinguishable performance of the staging and NM coordinates, for a given FP/HO diagonalization method. This is to be expected because both coordinates diagonalize the same energy function; i.e., the kinetic contribution of Eq. 2 for the FP method and the total for HO. The figure also shows that uncertainties in both  $E$  and  $C_V$  using the HO method are consistently smaller than those of the FP methods — about  $10\times$  and  $3\times$  smaller at the low and high temperatures, re-

spectively. Thus, for a given number of MC steps, the HO method generates more independent samples than the FP approach. This could be attributed the nature of both MC moves. While the HO method samples all degrees of freedom at once, including the centroid, the FP approach uses a separate translation move for the centroid. Therefore, the FP coordinates need more steps to reach to an independent sample. In addition, the trial potential associated with the collective move of the HO staging (Eq. 18) is different from that of FP staging (Eq. 19), which improves the acceptance rates for HO. The combined effect of these factors yield the lower uncertainties observed with the HO coordinates.

A closely related observation is that uncertainties in  $E$  and  $C_V$  using the FP methods increase substantially at low temperatures (below  $T \approx 0.2$ ), while the behavior of the HO methods is smooth and consistent in the whole temperature range. This could be understood in the context of the temperature dependence of the size of the ring-polymer. While the actual size at low temperatures (including  $T = 0$ ) is finite, the corresponding size of the free particle model (used in the FP methods) is given by the thermal de Broglie wavelength,<sup>25</sup> which diverges as the temperature approaches zero. Therefore, the proposed FP moves are not suitable at low temperatures, resulting in low acceptance rates and, hence, higher uncertainties. Different “fixes” are often adopted to enhance the sampling in the FP coordinates at low temperatures (e.g., partial FP sampling<sup>5</sup>). On the other hand, no such fixes are needed with the HO coordinates, which make them simpler to implement and, yet, more efficient.

### 3.2 PIMD

Here, we investigate the PIMD sampling efficiency based on the effect of time step size (accuracy) and statistical uncertainty (precision). Figure 4, shows the effect of time step on the energy (top) and heat capacity (bottom), for both HO (left) and AO (right) models. Results are shown for  $T = 0.5$ ; however, similar behavior is observed at all temperatures. To ensure nearly similar statistical uncertainties from all time steps, we fix the simulation time to  $t_{\text{sim}} = 10^7$ . We first recall that the fictitious masses used correspond to same oscillation ( $\omega$ ) of the centroid motion, such that comparing different methods at the same  $\Delta t$  is meaningful. For the HO model, the HO coordinates provide statistically exact solution, for both properties, with a maximum time step size corresponds to the Verlet threshold value ( $\Delta t = 2/\omega = 2.0$ ). This observation is already established analytically for the HO model, using the same BAOAB Verlet integrator.<sup>14</sup> On the other hand, substantially smaller values ( $\Delta t < 0.7$ ) are required for the FP method. This is a manifestation of the fact that the FP coordinates only diagonalize the FP contribution and, hence, they do not efficiently narrow the gap between multiple time scales. Although anharmonicity (right column) shows effects on the performance of all methods, results from HO coordinates is still the most efficient. A common observation for AO model, however, is that the heat capacity is apparently more sensitivity to  $\Delta t$  than the energy, for both HO and FP methods.

Moreover, for a given model and property, the size of the error bars from all methods appears to be statistically similar. To explain this, we first recognize the fact that the statistical uncertainty depends on the standard deviation (independent of  $\Delta t$ ) and the number of independent samples. For a given simulation time, the number of independent samples

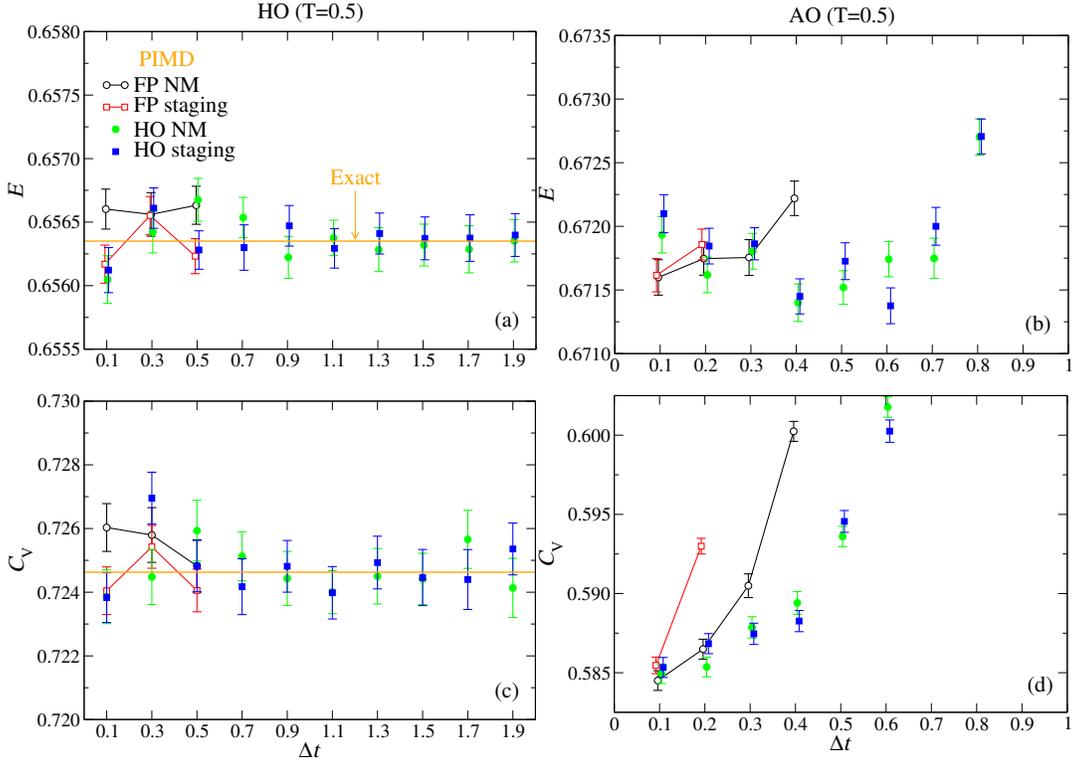


Figure 4: Dependence of the energy (top) and heat capacity (bottom) on the PIMD time step size, for both harmonic (left) and anharmonic (right) oscillators, at  $T = 0.5$ . Lines join the FP data points to guide the eye for the maximum time step size possible in each method. The horizontal orange lines represent the exact HO value (Eq. 39) at the given  $n = 40$  beads. The simulation time,  $t_{\text{sim}} = N_{\text{steps}}\Delta t = 10^7$  is kept fixed for all points. Data are shifted left/right for clarity.

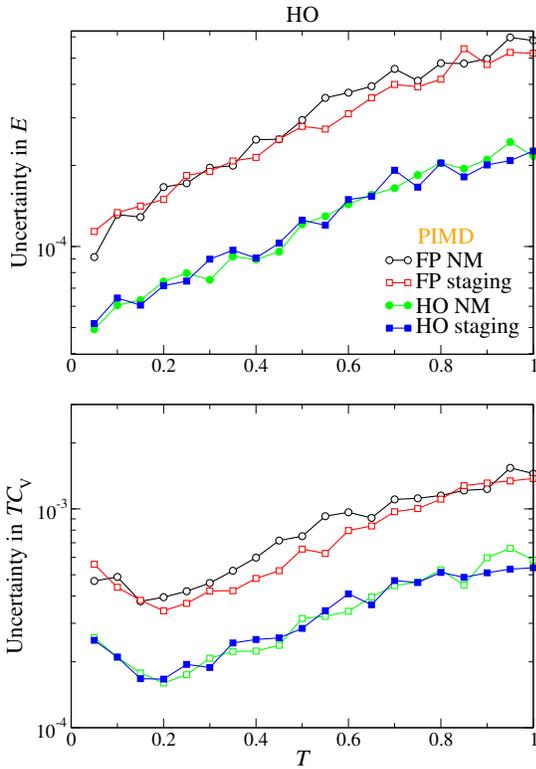


Figure 5: Same as Fig. 3, but using PIMD. The time step size used are  $\Delta t = 0.3$  and  $1.8$  for HO and FP coordinates, respectively.

depends on the slowest time scale present in the system, which corresponds here to the centroid mode. Therefore, since we force the centroid motion to match across all methods, the number of independent samples is expected to be similar, which yields the similar uncertainties we observe.

The direct effect of having large time step size, for a given  $t_{\text{sim}}$ , is to improve the precision in measured properties. Figure 5 depicts the temperature-dependence of the statistical uncertainty in estimating the energy (top) and heat capacity multiplied by temperature (bottom) of the HO model. Note that, the PIMD simulations ran using the largest time step sizes possible for *all* temperatures, without sacrificing the accuracy. This corresponds to  $\Delta t = 1.9$  and  $0.3$  for the HO and FP methods, respectively. The relative performance of the HO and FP methods is similar to that of the PIMC case, with the exception that the FP data do not diverge at low temperatures. This could be explained by the same aforementioned argument that the number of independent samples depends on the longest time scale, which is forced to match across all methods, regardless of the temperature. Hence, the number of independent samples is nearly constant, such that the variation of precision with temperature depends primarily on the standard deviation (a smooth function of  $T$ ).

### 3.3 Anharmonic Effects

Figure 6 depicts the effect of anharmonicity, in terms of the quartic spring constant  $k_4$  (Eq. 38b), on the precision of estimating the energy (top) and heat capacity (bottom), using PIMC (left) and PIMD (right) sampling methods. The data are shown for an intermediate temperature ( $T = 0.5$ ), how-

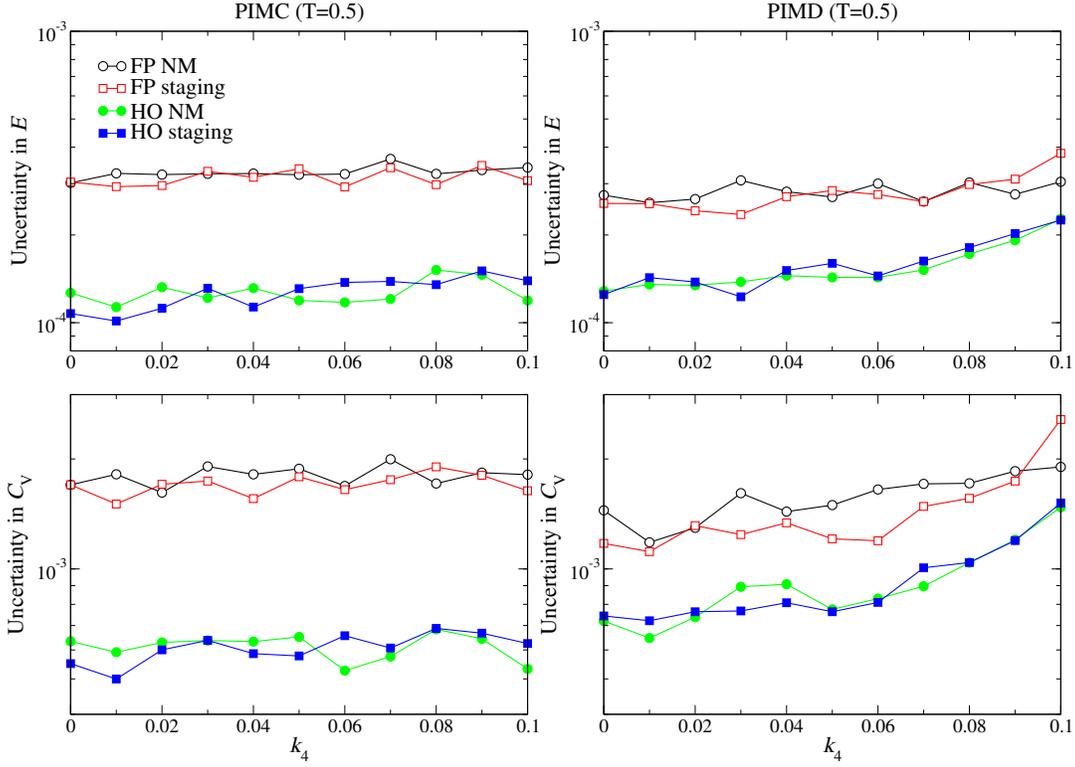


Figure 6: Dependence of the statistical uncertainty in the energy (top) and heat capacity (bottom) on the quartic anharmonic spring constant,  $k_4$ , using both PIMC (left) and PIMD (right). The maximum permissible  $\Delta t$  value is selected for each PIMD simulation to maintain accuracy.

ever a qualitatively similar behaviour is observed at other temperatures. For PIMC simulations, the relative performance of the HO and FP coordinates do not appear to be sensitive to anharmonicity, for both  $E$  and  $C_V$  properties. On the other hand, the PIMD data shows a marginal increase in uncertainties, with a slight reduction of the relative improvement of HO methods, as  $k_4$  increases. This could be partially understood using the fact that a smaller time step size is needed for the AO model, relative to the HO model (Fig. 4).

### 3.4 Computational Cost

Figure 7 represents the computational cost in terms of the CPU time needed for (simultaneously) computing  $E$  and  $C_V$  of the HO model as a function of the number of beads. We consider PIMC (top) and PIMD (bottom) simulations in both HO and FP coordinates, at  $T = 0.5$  (similar at other temperatures). The vertical dashed lines correspond to the number of beads that would be used according to the  $20\beta\hbar\omega$  criterion; i.e.,  $n = 200$  and  $20$  at  $T = 0.1$  and  $1.0$ , respectively. We notice that, the computational expense in the HO and FP methods (staging or NM) is nearly similar. This is because the only difference between these methods is the transformation formula. At large  $n$ , results from both PIMC and PIMD simulations show the expected  $\mathcal{O}(n)$  and  $\mathcal{O}(n^2)$  asymptotic scalings associated with staging and NM coordinates, respectively.

## 4 CONCLUSIONS

We have introduced novel staging coordinates that diagonalize the primitive path integral action of the harmonic

oscillator (HO) model. Hence, these coordinates are particularly suitable for systems with a harmonic character, such as quantum oscillators and crystalline systems. In this context, the method can be viewed as a natural extension to the standard staging coordinates, which only diagonalize the kinetic contribution of the action, corresponding to the free particle (FP) model. Additionally, the FP based normal mode (NM) coordinates were trivially extended to diagonalize the same HO action. In general, the staging coordinates are computationally preferable over the NM coordinates due its excellent  $\mathcal{O}(n)$  CPU scaling with the number of beads, rather than  $\mathcal{O}(n^2)$ . This becomes especially important at low temperatures, where large number of beads are required for statistical convergence.

We provide implementation schemes for applying the HO staging and NM coordinates in both PIMC and PIMD simulation methods. As an application, we use a one-dimensional HO model to assess the sampling efficiency using the new HO staging and NM coordinates, in comparison to the respective FP-based coordinates. The sampling performance was represented in terms of the statistical uncertainty in the total energy and heat capacity, using the centroid virial estimator. Moreover, the effect of anharmonicity on sampling was investigated using a one-dimensional harmonic+quartic anharmonic oscillator (AO).

Results from both PIMC and PIMD methods consistently show higher precision when using the HO coordinates, in comparison to the FP coordinates, in the whole temperature range considered. Moreover, the data suggest statistically indistinguishable performance from the staging and NM coordinates (for a given HO/FP method). This is to be expected because both coordinates diagonalize the same action (i.e., free particle/HO for the FP/HO methods).

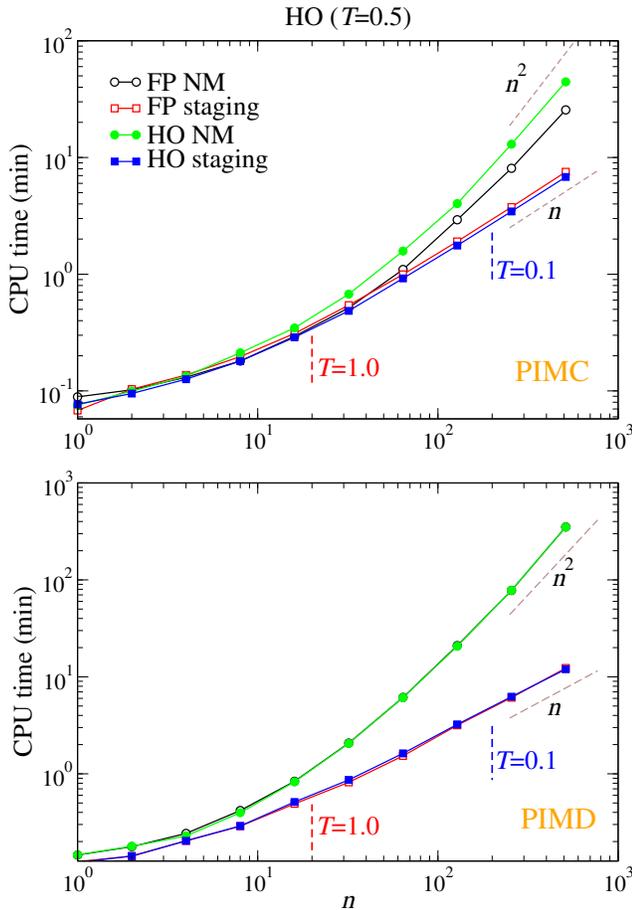


Figure 7: System size dependence of CPU time for (simultaneously) computing  $E$  and  $C_V$  of the AO model, at  $T = 0.5$ . Both staging and NM coordinates are used with PIMC (top) and PIMD (bottom) simulations. The vertical dashed lines correspond to the minimum  $n$  value ( $20\beta\hbar\omega$  here) required for statistical convergence at the given  $T$ .

In PIMC simulations, uncertainties in both energy and heat capacity using the FP coordinates substantially increase (diverge) at low temperatures, whereas the HO coordinates show a smooth behavior at all temperature. This FP behavior is caused by the diverging nature of the PI ring-polymer size as the temperature decreases, which results in low acceptance rate and, hence, large uncertainty. Moreover, anharmonic effects did not show influence of the absolute uncertainties from all methods.

On the other hand, the sampling performance of PIMD simulations was investigated in terms of the maximum allowable time step size  $\Delta t$ , without loss of accuracy, and the statistical uncertainty (after using these time steps). The HO model was perfectly sampled (i.e., no dependence on  $\Delta t$ ) using the HO coordinates. In this case, the allowed time step is only constraint by the maximum value for stable Verlet integrator, which is given by  $\Delta t = 2/\omega$ . However, much smaller time step sizes ( $\approx 20-30\%$  less) were tolerated by the FP coordinates. With anharmonicity, the maximum limits decrease for all method, yet the relative superior performance of the HO methods is still preserved. Moreover, using these maximum values, the precision in estimating both  $E$  and  $C_V$  using the HO coordinates is higher than that of the FP coordinates, at all temperatures considered. Unlike the PIMC case, both the absolute and relative uncertainties are affected by anharmonicity; but, with the HO coordinates continue to provide better relative precision.

Although the HO staging and NM coordinates are applied to quantum oscillator models, the framework is readily applicable to quantum crystals, using both PIMC and PIMD methods. In such systems, the Einstein crystal contribution (self force constants of the Hessian) acts as independent harmonic oscillators, such that same analysis adopted here can be applied. Of course, these realistic models have inherent anharmonic level, such that detailed studies on the performance of HO coordinates are needed. Moreover, although we only applied the HO coordinates to improve sampling, they can be used to modify the expressions of the PI estimators for an even more precise estimations. In fact, we are currently proceeding in this direction, using the harmonically mapped averaging (HMA) technique developed in our group.<sup>26</sup> With the simplicity of the HO coordinates and the no extra computational cost, these coordinates provide highly efficient alternative to traditional FP based coordinates when applied to systems with a harmonic character, such as quantum oscillators and crystals.

## Appendices

### A DERIVATION OF $x_i^*$ EXPRESSION

By definition (see Eqs. 3 and 4), the  $x_i^*$  position (for  $i > 0$ ) is obtained by minimizing  $V$  with respect to  $u_i$ , while fixing all other staging coordinates,

$$\frac{\partial V}{\partial u_i} = 0 \rightarrow x_i^*, \quad i > 0 \quad (41)$$

However, since  $V(\mathbf{x}(\mathbf{u}), \beta)$  is a natural function in the Cartesian coordinates (Eq. 2), we need to perform coordinate transformation in order to evaluate this derivative. Based on the definition of the HO staging transformation (Eq. 3), the Cartesian coordinate of bead  $i$  is given as  $x_i(u_0, \dots, u_i)$ . Using the chain rule, the derivative is then

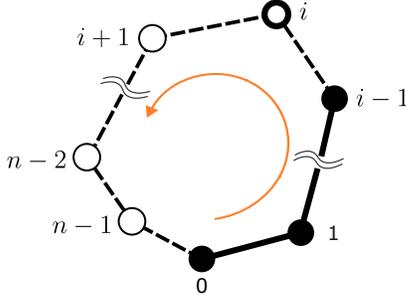


Figure 8: Illustration of the counterclockwise process to determine the instantaneous equilibrium position  $x_i^*$  of bead  $i$ , used with staging coordinates (Eq. 3). This is accomplished by minimizing the effective potential  $V$  with respect to the positions of beads  $i$  through  $n-1$  (open circles), while keeping the others fixed (filled circles). Lines represent PI harmonic interactions, with solid lines connecting the fixed beads, while dashed lines connecting the free beads. External interactions (not shown) are always assumed.

given as

$$\frac{\partial V}{\partial u_i} = \sum_{j=i}^{n-1} \frac{\partial V(\mathbf{x}, \beta)}{\partial x'_j} \frac{\partial x'_j}{\partial u_i} = 0, \quad i > 0 \quad (42)$$

where  $j = i, i+1, \dots, n-1$ . we assume the beads  $j = i$  to  $j = n-1$  to be “free” for the purpose of derivatives, while the other beads  $(0, \dots, i-1)$  are at their actual locations. Because Eq. 42 holds for any configuration and  $x'_j$  are independent variables, it is equivalent to the following set of conditions,

$$\frac{\partial V(\mathbf{x}, \beta)}{\partial x'_j} = 0, \quad j = i, \dots, n-1 \quad (43)$$

such that the  $x'_i$  solution is  $x_i^*$ . Therefore,  $x_i^*$  is defined as an instantaneous equilibrium position of bead  $i$ , such that  $V$  is minimum with respect to hypothetically “free” beads of indices  $i$  to  $n-1$ , while fixing the remaining beads at their actual locations.

Although we are interested in only the  $x'_i$  solution, we have to (simultaneously) satisfy all the  $n-i$  conditions given by Eq. 43, for each bead  $i$ . Using the real-space definition of  $V$  (Eq. 1b), these conditions reduce to the following system of  $n-i$  linear equations,

$$\begin{pmatrix} c & -1 & & & \\ -1 & c & -1 & & \\ & & \ddots & & \\ & & & -1 & c & -1 \\ & & & & -1 & c \end{pmatrix} \begin{pmatrix} x'_i \\ x'_{i+1} \\ \vdots \\ x'_{n-2} \\ x'_{n-1} \end{pmatrix} = \begin{pmatrix} x_{i-1} \\ 0 \\ \vdots \\ 0 \\ x_0 \end{pmatrix}, \quad i > 0 \quad (44)$$

where  $c = 2 + \epsilon^2$ , with  $\epsilon = \frac{\omega}{\omega_n} = \frac{\beta \hbar \omega}{n}$ . This coefficients matrix is symmetric and tridiagonal, with constant coefficients along the diagonal (i.e., Toeplitz). Fortunately, the inverse of this matrix is known analytically, with the following  $a, b$  components,<sup>27</sup>

$$S_{a,b} = \frac{\cosh((n-i+1-|a-b|)\alpha) - \cosh((n-i+1-a-b)\alpha)}{2 \sinh(\alpha) \sinh((n-i+1)\alpha)} \quad (45)$$

where  $a, b = 1, \dots, n-i$  and  $\alpha \equiv \cosh^{-1}(1 + \frac{1}{2}\epsilon^2)$ . The  $x'_i$

solution is then equal to  $x_i^*(x_0, x_{i-1}) = S_{1,1}x_{i-1} + S_{1,n-i}x_0$ , or

$$x_i^*(x_0, x_{i-1}) = \frac{\sinh(\alpha)x_0 + \sinh((n-i)\alpha)x_{i-1}}{\sinh((n-i+1)\alpha)}, \quad i > 0 \quad (46)$$

Accordingly, for our special PI harmonic action,  $x_i^*$  depends solely on two positions,  $x_0$  and  $x_{i-1}$ , of the fixed segment, which is a special case of the generalized definition (Eq. 3). This is not surprising since the intermediate fixed beads  $(1, \dots, i-2)$  do not interact directly with bead  $i$  (see Fig. 8).

## B DERIVATION OF $k_i$ EXPRESSION

Here we derive the expression for the HO staging spring constant  $k_i$ , through second derivatives of the PI effective potential of the HO model. Let us first express this potential in HO staging coordinates (Eq. 4) in terms of the Cartesian coordinates using Eq. 3,

$$V(\mathbf{x}, \beta) = \sum_{j=0}^{n-1} \frac{1}{2} \frac{m\omega_n^2}{n} (x_j - x_{j-1})^2 + \sum_{j=0}^{n-1} \frac{1}{2} \frac{m\omega^2}{n} x_j^2 = \sum_{j=0}^{n-1} \frac{1}{2} k_j [x_j - x_j^*(x_0, x_{j-1})]^2 \quad (47)$$

We will start by deriving  $k_i$  expression for  $i > 0$  and then the  $i = 0$  case. Taking the first derivative of both sides of Eq. 47, with respect to  $x_i$  (for  $i > 0$ ), yields

$$\frac{m\omega_n^2}{n} (2x_i - x_{i-1} - x_{i+1}) + \frac{m\omega^2}{n} x_i = k_i [x_i - x_i^*(x_0, x_{i-1})] - k_{i+1} [x_{i+1} - x_{i+1}^*(x_0, x_i)] \frac{\partial x_{i+1}^*(x_0, x_i)}{\partial x_i}, \quad i > 0 \quad (48)$$

Then, taking the second derivative, and recognizing that the second derivative of  $x_{i+1}^*$  with respect to  $x_i$  is zero (see Eq. 46), we get

$$k_i = 2 \frac{m\omega_n^2}{n} + \frac{m\omega^2}{n} - \left[ \frac{\partial x_{i+1}^*(x_0, x_i)}{\partial x_i} \right]^2 k_{i+1}, \quad i > 0. \quad (49)$$

According to Eq. 46, the derivative of  $x_i^*$  is

$$\frac{\partial x_{i+1}^*(x_0, x_i)}{\partial x_i} = \frac{\sinh((n-1-i)\alpha)}{\sinh((n-i)\alpha)}, \quad i > 0. \quad (50)$$

hence,

$$k_i = 2 \frac{m\omega_n^2}{n} + \frac{m\omega^2}{n} - \frac{\sinh^2((n-1-i)\alpha)}{\sinh^2((n-i)\alpha)} k_{i+1}, \quad i > 0$$

where, the spring constants  $k_i$  can be obtained in a backward recursion manner as

$$\begin{aligned}
k_{n-1} &= \frac{2m\omega_n^2}{n} + \frac{m\omega^2}{n} \\
k_{n-2} &= \frac{2m\omega_n^2}{n} + \frac{m\omega^2}{n} - \frac{\sinh^2(\alpha)}{\sinh^2(2\alpha)} k_{n-1} \\
k_{n-3} &= \frac{2m\omega_n^2}{n} + \frac{m\omega^2}{n} - \frac{\sinh^2(2\alpha)}{\sinh^2(3\alpha)} k_{n-2} \\
&\vdots \\
k_{n-i} &= \frac{2m\omega_n^2}{n} + \frac{m\omega^2}{n} - \frac{\sinh^2((n-1-i)\alpha)}{\sinh^2((n-i)\alpha)} k_{n-i+1} \\
&\vdots \\
k_1 &= \frac{2m\omega_n^2}{n} + \frac{m\omega^2}{n} - \frac{\sinh^2((n-2)\alpha)}{\sinh^2((n-1)\alpha)} k_2
\end{aligned}$$

Eventually, evaluating these expressions yields a simple closed form,

$$k_i = \frac{m\omega_n^2}{n} \frac{\sinh((n-i+1)\alpha)}{\sinh((n-i)\alpha)}, \quad i > 0. \quad (52)$$

Same procedure is followed to get  $k_0$ . Taking the first derivative of both sides of Eq. 47, with respect to  $x_0$ ,

$$\begin{aligned}
\frac{m\omega_n^2}{n} (2x_0 - x_{n-1} - x_1) + \frac{m\omega^2}{n} x_0 &= k_0 x_0 \\
- \sum_{i=1}^{n-1} [x_i - x_i^*(x_0, x_{i-1})] \frac{\partial x_i^*(x_0, x_{i-1})}{\partial x_0} k_i, & \quad (53)
\end{aligned}$$

then, taking the second derivative, and rearrange, we get

$$k_0 = \frac{2m\omega_n^2}{n} + \frac{m\omega^2}{n} - \sum_{i=1}^{n-1} \left[ \frac{\partial x_i^*(x_0, x_{i-1})}{\partial x_0} \right]^2 k_i. \quad (54)$$

According to Eq. 46, the derivative of  $x_i^*$  with respect to  $x_0$  is

$$\frac{\partial x_i^*(x_0, x_{i-1})}{\partial x_0} = \begin{cases} \frac{\sinh(\alpha) + \sinh((n-1)\alpha)}{\sinh(n\alpha)} & i = 1 \\ \frac{\sinh(\alpha)}{\sinh((n-i+1)\alpha)}, & i > 1 \end{cases} \quad (55)$$

and hence,

$$\begin{aligned}
k_0 &= \frac{2m\omega_n^2}{n} + \frac{m\omega^2}{n} - \left[ \frac{\sinh(\alpha) + \sinh((n-1)\alpha)}{\sinh(n\alpha)} \right]^2 k_1 \\
&- \sum_{i=2}^{n-1} \frac{\sinh^2(\alpha)}{\sinh^2((n-i+1)\alpha)} k_i. \quad (56)
\end{aligned}$$

Plugging  $k_i$  expressions (Eq. 52) into  $k_0$ , and simplify, yields a simple closed form,

$$k_0 = 2 \frac{m\omega_n^2}{n} \sinh(\alpha) \tanh\left(\frac{n\alpha}{2}\right) \quad (57)$$

## C VERIFICATION OF THE HO STAGING FORMULATION

Here we verify that the staging coordinates expression of the HO effective energy (Eq. 4) is equivalent to the Cartesian coordinates one (Eq. 2). Let us first plug the definition of  $x_i^*$  (Eq. 46) into the HO effective potential form (Eq. 4),

such that it is only a function of the Cartesian coordinates,

$$\begin{aligned}
V(\mathbf{x}, \beta) &= \frac{1}{2} k_0 x_0^2 \\
&+ \sum_{i=1}^{n-1} \frac{1}{2} k_i \left[ x_i - \frac{\sinh(\alpha) x_0 + \sinh((n-i)\alpha) x_{i-1}}{\sinh((n-i+1)\alpha)} \right]^2. \quad (58)
\end{aligned}$$

The task now is to show that this expression is equivalent to the original potential in Cartesian coordinates (Eq. 2). We achieve this by showing that both expressions have the same minimum (first derivative) and curvature (second derivative). Since both expressions are quadratic, higher order derivatives are zero in both cases. Therefore, same first and second derivatives is equivalent to same potential functions.

The configurations correspond to the minimum of the potential in Cartesian coordinates (2) are  $x_i = 0$ , with a zero potential. For the staging coordinates, the first derivative of Eq. 58, for  $i > 0$ , is

$$\begin{aligned}
0 &= \frac{\partial V(\mathbf{x}, \beta)}{\partial x_i} = k_i \left[ x_i - \frac{\sinh(\alpha) x_0 + \sinh((n-i)\alpha) x_{i-1}}{\sinh((n-i+1)\alpha)} \right] \\
&- k_{i+1} \frac{\sinh((n-i-1)\alpha)}{\sinh((n-i)\alpha)} \times \\
&\left[ x_{i+1} - \frac{\sinh(\alpha) x_0 + \sinh((n-i-1)\alpha) x_i}{\sinh((n-i)\alpha)} \right], \quad i > 0. \quad (59)
\end{aligned}$$

Similarly, the  $i = 0$  case is

$$\begin{aligned}
0 &= \frac{\partial V(\mathbf{x}, \beta)}{\partial x_0} = k_0 x_0 - k_1 \frac{\sinh(\alpha) + \sinh((n-1)\alpha)}{\sinh(n\alpha)} \times \\
&\left[ x_1 - \frac{\sinh(\alpha) + \sinh((n-1)\alpha)}{\sinh(n\alpha)} x_0 \right] \\
&- \sum_{j=2}^{n-1} k_j \frac{\sinh(\alpha)}{\sinh((n-j+1)\alpha)} \times \\
&\left[ x_j - \frac{\sinh(\alpha) x_0 + \sinh((n-j)\alpha) x_{j-1}}{\sinh((n-j+1)\alpha)} \right] \quad (60)
\end{aligned}$$

Since Eqs. 59 and 60 are valid for any  $x_i$ , these equations are satisfied if and only if all  $x_i = 0$ , which is the same minimum of the Cartesian space formulation.

Next, the second derivative of the Cartesian space potentials (Eq. 2) with respect to  $x_i$  is  $2 \frac{m\omega_n^2}{n} + \frac{m\omega^2}{n}$ . For staging, using Eq. 59, the second derivative for the  $i > 0$  case is

$$\begin{aligned}
\frac{\partial^2 V(\mathbf{x}, \beta)}{\partial x_i^2} &= k_i + k_{i+1} \frac{\sinh^2((n-i-1)\alpha) x_i}{\sinh^2((n-i)\alpha)} \\
&= \frac{m\omega_n^2}{n} \left[ \frac{\sinh((n-i+1)\alpha)}{\sinh((n-i)\alpha)} + \frac{\sinh((n-i-1)\alpha) x_i}{\sinh((n-i)\alpha)} \right] \\
&= 2 \frac{m\omega_n^2}{n} \cosh(\alpha) \\
&= 2 \frac{m\omega_n^2}{n} + \frac{m\omega^2}{n}, \quad i > 0 \quad (61)
\end{aligned}$$

which matches the second derivative of the Cartesian coordinates. For the  $i = 0$  case, taking the first derivative of

Eq. 60 with respect to  $x_i$  yields

$$\begin{aligned}
\frac{\partial^2 V(\mathbf{x}, \beta)}{\partial x_0^2} &= k_0 + k_1 \left[ \frac{\sinh(\alpha) + \sinh((n-1)\alpha)}{\sinh(n\alpha)} \right]^2 \\
&+ \sum_{j=1}^{n-1} k_j \left[ \frac{\sinh(\alpha)}{\sinh((n-j+1)\alpha)} \right]^2 \\
&= 2 \frac{m\omega_n^2}{n} \sinh(\alpha) \tanh\left(\frac{n\alpha}{2}\right) \\
&+ \frac{m\omega_n^2}{n} \frac{(\sinh(\alpha) + \sinh((n-1)\alpha))^2}{\sinh((n-1)\alpha) \sinh(n\alpha)} \\
&+ \frac{m\omega_n^2}{n} \sum_{j=2}^{n-1} \frac{\sinh^2(\alpha)}{\sinh((n-j)\alpha) \sinh((n-j+1)\alpha)} \\
&= 2 \frac{m\omega_n^2}{n} + \frac{m\omega^2}{n}, \tag{62}
\end{aligned}$$

where the last line is obtained from Mathematica.<sup>28</sup> This completes the verification that the staging expression (Eq. 58) is equal to the Cartesian space one (Eq. 2).

## D CLOCKWISE VERSION OF STAGING

The staging expressions derived in this work are based on a forward  $(0, 1, \dots, n-1)$  counterclockwise representation (Fig. 8). However, the traditional FP staging known in literature is based on a backward  $(0, n-1, n-2, \dots, 1)$  clockwise formulation. Clearly, there is no difference between both versions since only beads labelling is different. However, to show equivalence between both FP staging versions and to provide an alternative option for HO staging, which might be of more convenience to some, we present here the clockwise version of both HO and FP stagings. This achieved by simply replacing the  $x_{i-1}$  position by  $x_{i+1}$  and the  $n-i$  index by  $i$  (not including positions subscripts), which results in the following HO and FP staging expressions.

### Clockwise HO staging

The clockwise version of the HO staging coordinates (Eq. 7) is given by a non-recursive form,

$$u_i = x_i - \frac{\sinh(\alpha) x_0 + \sinh(i\alpha) x_{i+1}}{\sinh((i+1)\alpha)}, \quad i > 0 \tag{63}$$

where  $u_0 = x_0$  and  $x_n = x_0$ . Accordingly the inverse transformation ( $\mathbf{x} \leftarrow \mathbf{u}$ ) is given by the following recursive (from  $i = n-1$  to 1) relation:

$$x_i = u_i + \frac{\sinh(\alpha) u_0 + \sinh(i\alpha) x_{i+1}}{\sinh((i+1)\alpha)}, \quad i > 0 \tag{64}$$

Similarly, the clockwise version of the staging spring constants (Eq. 6) are

$$k_i = \frac{m\omega_n^2}{n} \begin{cases} 2 \sinh(\alpha) \tanh\left(\frac{n\alpha}{2}\right), & i = 0 \\ \frac{\sinh((i+1)\alpha)}{\sinh(i\alpha)}, & i > 0 \end{cases} \tag{65}$$

### Clockwise FP staging

For FP staging, the clockwise version of the staging transformation (Eq. 9) is given by

$$u_i = x_i - \frac{x_0 + ix_{i+1}}{i+1}, \quad i > 0 \tag{66}$$

where  $u_0 = x_0$  and  $x_n = x_0$ . From this, the inverse transformation ( $\mathbf{x} \leftarrow \mathbf{u}$ ) is given by the following recursive (from  $i = n-1$  to 1) relation:

$$x_i = u_i + \frac{u_0 + ix_{i+1}}{i+1}, \quad i > 0 \tag{67}$$

Similarly, the clockwise version of the FP staging spring constant is given from Eq. 11 as,

$$k_i = \begin{cases} 0, & i = 0 \\ \frac{m\omega_n^2}{n} \frac{i+1}{i}, & i > 0 \end{cases} \tag{68}$$

## E FINITE SIZE EFFECTS

Figure 9 shows the finite size effects (FSE) for the energy (top) and heat capacity (bottom) of the HO model, at  $T = 0.5$ . We choose the HO model in order to verify the accuracy of our estimations against the exact analytical values of the model. However, similar results are observed with the AO model (not shown). Since the leading FSE term is known to be  $(\beta\hbar\omega/n)^2$ , we represent the data in terms of  $1/n^2$ . Within the number of beads range considered ( $n = 5$  to 40), it is evident that the trend of both  $E$  and  $C_V$  is linear, which is consistent with the leading term behavior. The orange lines represent the analytical values (Eq. 39), which clearly show the accuracy of our estimates at all system sizes considered.

Moreover, as mentioned earlier, we use number of beads equal to  $n = 20\beta\hbar\omega$ , to ensure converged results. To verify this, we consider this specific value in Fig. 9 (leftmost data point). It is clear that the energy and heat capacity data at that point are statistically consistent with the continuum limit (Eq. 40). Similar observations were reported (not shown) using other temperatures and HO frequencies, for both HO and AO models.

Figure 10 shows the effect of system size on the precision of estimating the energy of HO model, using PIMC (top) and PIMD (bottom) simulations. We present both HO and FP based coordinates, at low ( $T = 0.1$ , dashed lines) and high ( $T = 1.0$ , solid lines) temperatures. We do not find a noticeable difference in the behavior between the HO and FP methods. In addition, the behavior beyond the converged number of beads ( $n = 20\beta\hbar\omega$ , represented by vertical lines) is flat, which is a well established observation in the literature.

## ACKNOWLEDGMENTS

Computational resources were provided by the Center for Computational Research (CCR), University at Buffalo, NY.

## References

- (1) Markland, T. E.; Ceriotti, M. Nuclear quantum effects enter the mainstream. *Nature Reviews Chemistry* **2018**, *2*, 0109.

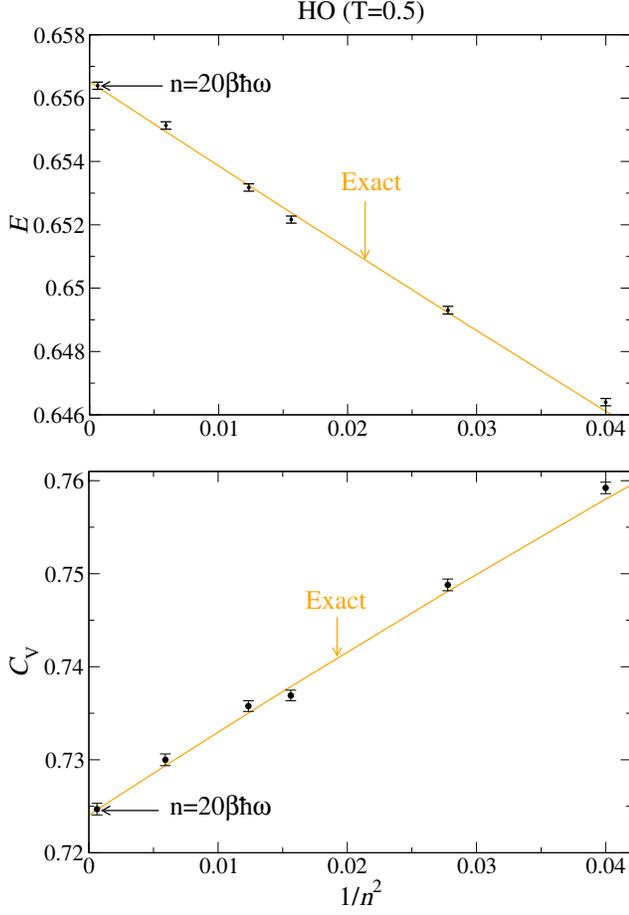


Figure 9: Dependence of the energy (top) and heat capacity (bottom) on the number of beads for the HO model, at  $T = 0.5$ . The orange lines represent the exact values (Eq. 39). The leftmost point ( $n = 40$ ) corresponds to our  $20\beta\hbar\omega$  criterion for statistical convergence. The data were generated using HO staging coordinates in PIMC.

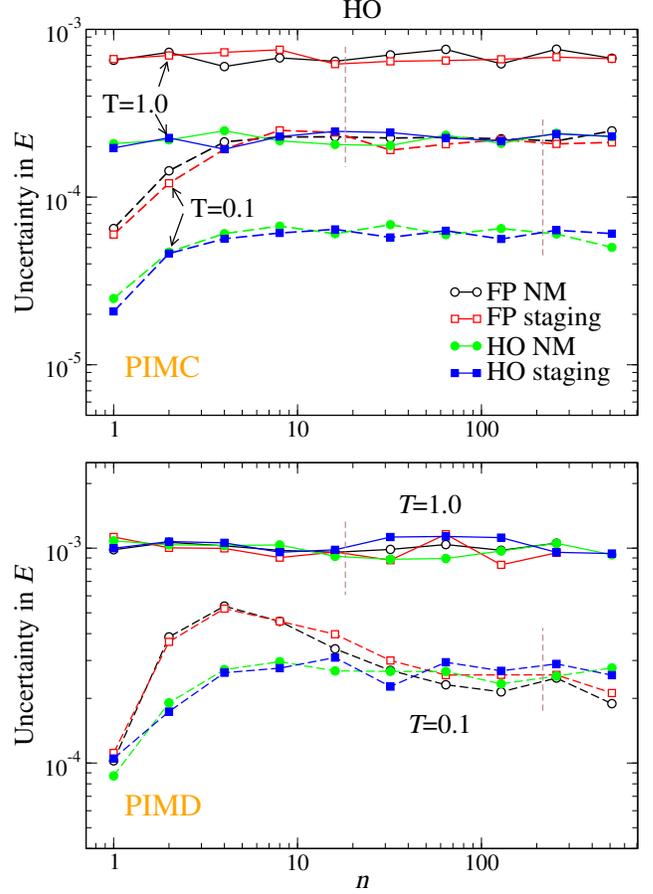


Figure 10: Dependence of the uncertainty in energy on the number of beads,  $n$ , for the HO model, using PIMC (top) and PIMD (bottom) at low ( $T = 0.1$ ) and high ( $T = 1.0$ ) temperatures. The vertical lines correspond to the  $n = 20\beta\hbar\omega$  criterion for statistical convergence; see Fig. 9.

- (2) Herrero, C. P.; Ramirez, R. Path-integral simulation of solids. *Journal of Physics: Condensed Matter* **2014**, *26*, 233201.
- (3) Feynman, R. P.; Hibbs, A. R. *Quantum mechanics and path integrals*; International series in pure and applied physics; McGraw-Hill: New York, NY, 1965.
- (4) Tuckerman, M. E. *Statistical mechanics: theory and molecular simulation*; Oxford university press, 2023.
- (5) Tuckerman, M. E.; Berne, B. J.; Martyna, G. J.; Klein, M. L. Efficient molecular dynamics and hybrid Monte Carlo algorithms for path integrals. *The Journal of Chemical Physics* **1993**, *99*, 2796–2808.
- (6) Hall, R. W.; Berne, B. J. Nonergodicity in path integral molecular dynamics. *The Journal of chemical physics* **1984**, *81*, 3641–3643.
- (7) Tuckerman, M. E.; Marx, D.; Klein, M. L.; Parrinello, M. Efficient and general algorithms for path integral Car–Parrinello molecular dynamics. *The Journal of chemical physics* **1996**, *104*, 5579–5588.
- (8) Herman, M.; Bruskin, E.; Berne, B. On path integral Monte Carlo simulations. *The Journal of Chemical Physics* **1982**, *76*, 5150–5155.
- (9) Cao, J.; Berne, B. A Born–Oppenheimer approximation for path integrals with an application to electron solvation in polarizable fluids. *The Journal of chemical physics* **1993**, *99*, 2902–2916.
- (10) Friesner, R. A.; Levy, R. M. An optimized harmonic reference system for the evaluation of discretized path integrals. *The Journal of chemical physics* **1984**, *80*, 4488–4495.
- (11) Whitfield, T. W.; Martyna, G. J. Low variance energy estimators for systems of quantum Drude oscillators: Treating harmonic path integrals with large separations of time scales. *The Journal of chemical physics* **2007**, *126*.
- (12) Li, D.; Voth, G. A. A path integral Einstein model for characterizing the equilibrium states of low temperature solids. *The Journal of chemical physics* **1992**, *96*, 5340–5353.
- (13) Markland, T. E.; Manolopoulos, D. E. An efficient ring polymer contraction scheme for imaginary time path integral simulations. *The Journal of chemical physics* **2008**, *129*, 024105.
- (14) Leimkuhler, B.; Matthews, C. Robust and efficient configurational molecular sampling via Langevin dynamics. *The Journal of Chemical Physics* **2013**, *138*, 174102.
- (15) Allen, M.; Tildesley, D. *Computer Simulation of Liquids*; Oxford science publications; Oxford University Press, 2017.
- (16) Liu, J.; Li, D.; Liu, X. A simple and accurate algorithm for path integral molecular dynamics with the Langevin thermostat. *The Journal of chemical physics* **2016**, *145*.
- (17) Müser, M. On new efficient algorithms for PIMC and PIMD. *Computer physics communications* **2002**, *147*, 83–86.
- (18) Witt, A.; Ivanov, S. D.; Shiga, M.; Forbert, H.; Marx, D. On the applicability of centroid and ring polymer path integral molecular dynamics for vibrational spectroscopy. *The Journal of chemical physics* **2009**, *130*.
- (19) Shiga, M.; Shinoda, W. Calculation of heat capacities of light and heavy water by path-integral molecular dynamics. *The Journal of chemical physics* **2005**, *123*.
- (20) Frenkel, D.; Smit, B. *Understanding molecular simulation: from algorithms to applications*; Elsevier, 2023.
- (21) Schweizer, K. S.; Stratt, R. M.; Chandler, D.; Wolynes, P. G. Convenient and accurate discretized path integral methods for equilibrium quantum mechanical calculations. *The Journal of Chemical Physics* **1981**, *75*, 1347–1364.
- (22) Chin, S. A. Analytical evaluations of the path integral Monte Carlo thermodynamic and Hamiltonian energies for the harmonic oscillator. *The Journal of Chemical Physics* **2023**, *159*.
- (23) Liu, J. Path integral Liouville dynamics for thermal equilibrium systems. *The Journal of chemical physics* **2014**, *140*.
- (24) Schultz, A. J.; Kofke, D. A. Etomica: An object-oriented framework for molecular simulation. 2015.
- (25) Ceperley, D. M. Path integrals in the theory of condensed helium. *Reviews of Modern Physics* **1995**, *67*, 279.
- (26) Schultz, A. J.; Moustafa, S. G.; Lin, W.; Weinstein, S. J.; Kofke, D. A. Reformulation of ensemble averages via coordinate mapping. *Journal of chemical theory and computation* **2016**, *12*, 1491–1498.
- (27) Hu, G.; O’Connell, R. F. Analytical inversion of symmetric tridiagonal matrices. *Journal of Physics A: Mathematical and General* **1996**, *29*, 1511.
- (28) Inc., W. R. Mathematica, Version 14.0. <https://www.wolfram.com/mathematica>, Champaign, IL, 2024.