

# Probing the partition function for temperature-dependent potentials with nested sampling

Lune Maillard,<sup>1, a)</sup> Philippe Depondt,<sup>1</sup> Fabio Finocchi,<sup>1</sup> Simon Huppert,<sup>1</sup> Thomas Plé,<sup>2</sup> Julien Salomon,<sup>3</sup> and Martino Trassinelli<sup>1</sup>

<sup>1)</sup>*Sorbonne Université, CNRS, Institut des Nanosciences de Paris, INSP, F-75005 Paris, France*

<sup>2)</sup>*Sorbonne Université, CNRS, Laboratoire de Chimie Théorique, LCT, F-75005 Paris, France*

<sup>3)</sup>*Laboratoire Jacques-Louis Lions, Sorbonne Université and ANGE, INRIA Paris, France*

(Dated: 4 December 2025)

Thermodynamic properties can be in principle derived from the partition function, which, in many-atom systems, is hard to evaluate as it involves a sum on the accessible microscopic states. Recently, the partition function has been computed via nested sampling, relying on Bayesian statistics, which is able to provide the density of states as a function of the energy in a single run, independently of the temperature. This appealing property is lost whenever the potential energy that appears in the partition function is temperature-dependent: for instance, mean-field effective potential energies or the quantum partition function in the path-integral formalism. For these cases, the nested sampling must be carried out at each temperature, which results in a massive increase of computational time. Here, we introduce and implement a new method, that is based on an extended partition function where the temperature is considered as an additional parameter to be sampled. The extended partition function can be evaluated by nested sampling in a single run, so to restore this highly desirable property even for temperature-dependent effective potential energies. We apply this original method to compute the quantum partition function for harmonic potentials and Lennard-Jones clusters at low temperatures and show that it outperforms the straightforward application of nested sampling for each temperature within several temperature ranges.

## I. INTRODUCTION

The central quantity in statistical mechanics of many-atom systems is the partition function, from which all the thermodynamic properties can in principle be derived. For a system of  $N$  classical particles, the partition function  $Z$  is usually expressed as a multidimensional integral over  $6N$  degrees of freedom, corresponding to the position vector  $\mathbf{x}$  and the associated momenta  $\mathbf{p}$  of the  $N$  particles:

$$Z(\beta) \propto \int d\mathbf{x} d\mathbf{p} e^{-\beta H(\mathbf{x}, \mathbf{p})} \quad (1)$$

with  $\beta$  the inverse temperature and  $H(\mathbf{x}, \mathbf{p})$  the Hamiltonian. Furthermore, via the statistical weight  $\exp(-\beta H)$ ,  $Z$  is a function of the temperature. In practice, the computation of thermodynamic properties is very difficult as it implies sampling the high dimensional configurational space, which is computationally very demanding except for some very simple systems, subject for instance to a harmonic potential.

There is therefore a need for efficient numerical methods to tackle the problem of efficiently sampling the configurational space of a multi-atomic system. In the past, a variety of methods have been developed: parallel tempering<sup>1–3</sup>, simulated annealing<sup>4–6</sup> and metadynamics<sup>7,8</sup> among others. In this work, we use the nested sampling algorithm<sup>9,10</sup>. This algorithm, initially developed for data analysis applications, is a very general method that can be used to minimize the distance between models and experimental data: the model's pa-

rameter space is explored with an increasing likelihood constraint while simultaneously estimating the volume of parameter space in which the likelihood is above the constraint. This allows to better compare the relevance of different models, which standard least-square fit methods have difficulty doing. Bayesian model selection based on nested sampling has been applied in a variety of fields such as cosmology<sup>11</sup>, biology<sup>12</sup>, microseismic events<sup>13</sup> and many other fields, and its convergence has formally been proven<sup>14,15</sup>.

In a different context, nested sampling can be used to explore the phase space of a physical system with a lowering energy constraint while simultaneously estimating the density of states (DOS), i.e., the number of states available at a given energy: the partition function and its derivatives can therefore be estimated in a very elegant manner. It has been applied to study the Potts model<sup>16</sup>, Lennard-Jones cluster<sup>17,18</sup>, transition path sampling<sup>19</sup>, and the phase space of zirconium<sup>20</sup> and water clusters<sup>21</sup>, among others. One of the advantage of this algorithm is that, for temperature-independent potentials, the thermodynamic quantities can be computed at all temperatures with only a single nested sampling exploration<sup>17</sup>, contrary to the methods mentioned above.

However, this advantage is lost when the potential depends on the temperature, which is, for example, the case of mean-field theories<sup>22,23</sup> where a physical variable is replaced by its temperature-dependent average<sup>23</sup>, the Debye model<sup>24</sup> used for computing the specific heat of solids<sup>25</sup>, or the path-integral formalism<sup>26,27</sup> which takes into account the quantum nature of the nuclei. In this work, we consider and compare two methods to use nested sampling on such potentials: the *direct method*, used in Ref. 21, which consists in performing several explorations in a pre-determined temperature range, one

<sup>a)</sup>lune.maillard@insp.upmc.fr

for each temperature that we wish to sample, and the *extended partition function method* that allows to perform a single exploration in the case of a temperature-dependent potential and thus restores the "all-at-once" character of nested sampling.

A particularly interesting field of application of these two methods is the path-integral formalism which is used to take into account nuclear quantum effects (NQEs) that can have a significant impact on the structure and dynamics of systems containing light element such as hydrogen<sup>27</sup>. The most direct manifestation of NQEs are isotope effects: indeed, for a classical system described by Eq. (1), the thermodynamic properties do not depend on the nuclear mass. Actually, isotope substitution can affect the properties of common compounds such as the melting temperature of water that increases by 4 K when replacing hydrogen with deuterium<sup>28,29</sup>. In some cases, isotope substitution is known to even suppress some phase transitions entirely as in some ferroelectric materials<sup>30,31</sup>. NQEs are also often relevant to phase transitions in high-pressure materials<sup>32–34</sup>. To account for these NQEs in simulations in a numerically feasible manner, the path-integral formalism<sup>26</sup> consists in replacing the quantum nuclei by classical polymers made of  $P$  replicas to mimic the quantum delocalization. This increases the numbers of degrees of freedom by a factor  $P$ , where  $P$  is typically of the order of 10–100, making the computation of the thermodynamic quantities, such as the DOS, even more complex. In an effort to test new methods that can overcome this bottleneck, we test the two approaches based on nested sampling (the direct method and the extended partition function method) on two specific examples: the harmonic potential, for which analytical values are known, and atoms interacting via a Lennard-Jones potential. Another method that can be used to study isotope effects is thermodynamic integration<sup>35,36</sup> where the derivative of the free energy is integrated over the mass.

We first present the nested sampling algorithm in Section II for the cases of temperature-independent and of temperature-dependent potentials for which we compare the direct and extended partition function methods. Next, in Section III, we recall the path-integral formalism. We then consider two applications: the quantum harmonic potential in Section IV and Lennard-Jones clusters in Section V. Finally, we conclude and give some perspectives for future work in Section VI.

## II. NESTED SAMPLING AND THE PARTITION FUNCTION

### A. Temperature-independent potential - General equations

Let us consider a system composed of  $N$  nuclei interacting via a potential  $V(\mathbf{x})$ , which depends only on the positions of the nuclei  $\mathbf{x}$  and neither on their momenta nor on the temperature. For this system, the partition function  $Z$  can be written as

$$Z(\beta) = \frac{1}{N! h^{3N}} \int d\mathbf{x} d\mathbf{p} \exp \left( -\beta \left( \frac{1}{2} \sum_{i=1}^{3N} \frac{p_i^2}{m_i} + V(\mathbf{x}) \right) \right) \quad (2)$$

where  $\mathbf{p} = (p_1, \dots, p_{3N})$  is the vector of momenta,  $(m_1, \dots, m_{3N})$  is the vector of masses,  $\beta = \frac{1}{k_B T}$  the inverse temperature and  $h$  the Planck constant. In this work, we consider the case where  $m_i = m$  for all  $1 \leq i \leq 3N$  to simplify the notations but the generalization to different masses is straightforward. In this case, the partition function can be broken-down into two terms: a momentum-dependent one (kinetic contribution)

$$Z_k(\beta) = \frac{1}{N! h^{3N}} \int d\mathbf{p} \exp \left( -\beta \left( \frac{1}{2} \sum_{i=1}^{3N} \frac{p_i^2}{m} \right) \right) = \left( \sqrt{\frac{2\pi m}{(N!)^2 \beta h^2}} \right)^{3N} \quad (3)$$

and a position-dependent one (configurational contribution)<sup>37</sup>

$$Z_c(\beta) = \int d\mathbf{x} e^{-\beta V(\mathbf{x})} = \int dE \rho(E) e^{-\beta E}, \quad (4)$$

where  $\rho(E)$  is the density of states, that is the number of microscopic configurations with energies between  $E$  and  $E+dE$ :

$$\rho(E) = \int d\mathbf{x} \delta(E - V(\mathbf{x})), \quad (5)$$

and  $\delta(\cdot)$  is the Dirac operator. The partition function  $Z$  becomes the product of two contributions

$$Z(\beta) = Z_k(\beta) \times Z_c(\beta). \quad (6)$$

The relevant thermodynamic properties of the system can be deduced from  $Z(\beta)$ . As an example, the internal energy can be computed as

$$U = -\frac{\partial \log(Z)}{\partial \beta} = \frac{3N}{2} k_B T + \frac{\int dE \rho(E) E e^{-\beta E}}{Z_c} = U_k + U_c, \quad (7)$$

where  $U_k = 3Nk_B T/2$  and  $U_c = (\int dE \rho(E) E e^{-\beta E})/Z_c$  are the momentum and position contributions, i.e., the average kinetic and potential energies, respectively. The heat capacity can be computed as

$$C_v = \frac{\partial U}{\partial T} = \frac{3N}{2} k_B + \left( \frac{\int dE \rho(E) E^2 e^{-\beta E}}{k_B T^2 Z_c} - \frac{(\int dE \rho(E) E e^{-\beta E})^2}{k_B T^2 Z_c^2} \right) = C_{v,k} + C_{v,c}, \quad (8)$$

where  $C_{v,k} = 3Nk_B/2$  and

$$C_{v,c} = \frac{\int dE \rho(E) E^2 e^{-\beta E}}{k_B T^2 Z_c} - \frac{(\int dE \rho(E) E e^{-\beta E})^2}{k_B T^2 Z_c^2} \quad (9)$$

are the momentum and position contributions respectively.

In general, the position-dependent term  $Z_c$  does not have an analytical form. Thermodynamic properties such as  $U$  and  $C_v$  thus have to be estimated numerically, generally using importance sampling of the phase space, for instance with some flavor of the Monte-Carlo Metropolis algorithm<sup>38</sup>. Here, we use nested sampling, a method that has its origins in Bayes' theory of probabilities<sup>9</sup>.

## B. The nested sampling algorithm and estimation of the partition function

Let us recall the main steps of the nested sampling algorithm<sup>9,17</sup> to sample the partition function:

1. A relevant sampling space for positions is defined and  $K$  points, called *live points*, are uniformly sampled from this space. The integrals in Eqs. (4) and (6) to (9) will then be replaced by discrete sums over the sampling points.
2. At each iteration  $\ell$ , the point  $\mathbf{x}_{\text{old}}$  associated with the highest energy is removed and replaced by a random point  $\mathbf{x}_{\text{new}}$  with an energy that is strictly lower:  $V(\mathbf{x}_{\text{new}}) < V(\mathbf{x}_{\text{old}})$ . The partition function can then be estimated by the trapezoid rule<sup>10</sup>:

$$Z_c(\beta) \approx \sum_{i=1}^{\ell} w_i e^{-\beta E_i}, \quad (10)$$

where  $E_\ell = V(\mathbf{x}_{\text{old}})$  is the energy of the removed point at iteration  $\ell$  and

$$w_\ell = \frac{1}{2} \left( e^{-\frac{\ell-1}{K}} - e^{-\frac{\ell+1}{K}} \right) \quad (11)$$

is an approximation of the DOS:  $w_\ell \approx \rho(E)$ . Equation (10) is therefore a sum over the points removed at all previous iterations ( $E_i$  and  $w_i$  are respectively the energy of the removed points and the approximation of the DOS at iteration  $i$ , for  $i \leq \ell$ ). More details are given in Appendix A.

3. This procedure is repeated until the current contribution is small compared to previous contributions:  $\log(w_\ell e^{-\beta_s E_\ell}) - \log(c_{\max}) < \delta$  with  $c_{\max} = \max_{i \leq \ell} (w_i e^{-\beta_s E_i})$  at an inverse temperature  $\beta_s$  chosen by the user. This temperature  $T_s = 1/(k_B \beta_s)$  needs to be small to ensure a good exploration of the phase space: it has to be inferior to the lowest temperature studied<sup>39</sup>. Here, we use  $\delta = -10$ , a value that satisfies our requirements.

At the end of the procedure, we thus have a collection of  $(w_i, E_i)$  values, i.e., the DOS estimate  $w_i$  at the energy  $E_i$ , from which we can compute an approximated partition function using Eq. (10) with  $\ell = n_{\text{iter}}$ , the number of iterations. We therefore sum over all the states that have been removed during the exploration, that is, until an energy minimum is reached. In the following, for simplicity, we will omit the bounds of the sum.

The advantage when using nested sampling on temperature-independent potentials is that  $Z$  can be computed at all temperatures with only one exploration via nested sampling<sup>17</sup> since the series of sampled  $E_i$  does not depend on temperature but on the distribution of states (that is, the DOS) which is independent of  $\beta$  as the potential  $V(\mathbf{x})$  (see Eqs (4) and (5)). This is also the case for the potential energy contribution to the

internal energy and the heat capacity (see Eqs. (7) and (8)), which can be recovered as<sup>17</sup>

$$U_c(\beta) = \frac{\sum_i w_i E_i e^{-\beta E_i}}{Z_c(\beta)} \quad (12)$$

and

$$C_{V,c}(\beta) = \frac{\sum_i w_i E_i^2 e^{-\beta E_i}}{k_B T^2 Z_c(\beta)} - \frac{(\sum_i w_i E_i e^{-\beta E_i})^2}{k_B T^2 Z_c(\beta)^2}, \quad (13)$$

respectively.

All along the present work, we use the `nested_fit` program<sup>39-42</sup> which can be used both for model comparison in data analysis and for the estimation of partition functions in materials science. `Nested_fit` has previously been used to study systems in which the potential energy does not depend on temperature, namely the harmonic potential<sup>39</sup> and Lennard-Jones clusters<sup>43</sup>. In both cases, the program was able to recover the correct curves for the heat capacity. Here, we aim at studying the more complex case of temperature-dependent potentials, with a particular focus on quantum systems described within the path-integral framework.

## C. Temperature-dependent potential

By temperature-dependent potential, we mean cases in which the partition function is written

$$Z(\beta) = Z_k(\beta) \times \int d\mathbf{x} e^{-\beta V(\mathbf{x}, \beta)}, \quad (14)$$

where  $Z_k$  is the position-independent prefactor, that can be computed separately, and  $V(\mathbf{x}, \beta)$  is the effective potential depending on the positions  $\mathbf{x}$  and inverse temperature  $\beta$  (but not on the momenta). In analogy to Eq. (4), we define

$$Z_c(\beta) = \int d\mathbf{x} e^{-\beta V(\mathbf{x}, \beta)} \quad (15)$$

the position-dependent part of the partition function. In particular, from Eq. (5), it follows that the DOS depends now on  $T$ :

$$\rho(E, \beta) = \int d\mathbf{x} \delta(E - V(\mathbf{x}, \beta)). \quad (16)$$

Similarly to Eqs. (7) and (8), the internal energy and heat capacity can be recovered from the partition function. However, since the potential now depends on the temperature, the derivatives with respect to  $\beta$  contain additional terms and the expressions are more involved:

$$U_c = \frac{\int d\mathbf{x} \left( V(\mathbf{x}, \beta) + \beta \frac{\partial V(\mathbf{x}, \beta)}{\partial \beta} \right) e^{-\beta V(\mathbf{x}, \beta)}}{Z_c} \\ = \left\langle V(\mathbf{x}, \beta) + \beta \frac{\partial V(\mathbf{x}, \beta)}{\partial \beta} \right\rangle, \quad (17)$$

$$\begin{aligned}
C_{V,c} &= \frac{\int d\mathbf{x} \left( V(\mathbf{x}, \beta) + \beta \frac{\partial V(\mathbf{x}, \beta)}{\partial \beta} \right)^2 e^{-\beta V(\mathbf{x}, \beta)}}{k_B T^2 Z_c} \\
&\quad - \frac{\left( \int d\mathbf{x} \left( V(\mathbf{x}, \beta) + \beta \frac{\partial V(\mathbf{x}, \beta)}{\partial \beta} \right) e^{-\beta V(\mathbf{x}, \beta)} \right)^2}{k_B T^2 Z_c^2} \\
&\quad - \frac{\int d\mathbf{x} \left( 2 \frac{\partial V(\mathbf{x}, \beta)}{\partial \beta} + \beta \frac{\partial^2 V(\mathbf{x}, \beta)}{\partial \beta^2} \right) e^{-\beta V(\mathbf{x}, \beta)}}{k_B T^2 Z_c} \\
&= \frac{\left\langle \left( V(\mathbf{x}, \beta) + \beta \frac{\partial V(\mathbf{x}, \beta)}{\partial \beta} \right)^2 \right\rangle - \left\langle V(\mathbf{x}, \beta) + \beta \frac{\partial V(\mathbf{x}, \beta)}{\partial \beta} \right\rangle^2}{k_B T^2} \\
&\quad - \frac{\left\langle 2 \frac{\partial V(\mathbf{x}, \beta)}{\partial \beta} + \beta \frac{\partial^2 V(\mathbf{x}, \beta)}{\partial \beta^2} \right\rangle}{k_B T^2}, \tag{18}
\end{aligned}$$

where  $\langle \cdot \rangle$  denotes the average on the position coordinate  $\mathbf{x}$  weighted by the probability distribution  $\exp(-\beta V(\mathbf{x}, \beta))/Z_c$ . We see that the first and second derivatives of the potential  $V$  with respect to  $\beta$  appear in the above expressions. Equations (7) and (8) for temperature-independent potentials can be recovered by taking  $(\partial V(\mathbf{x}, \beta))/(\partial \beta) = (\partial^2 V(\mathbf{x}, \beta))/(\partial \beta^2) = 0$ .

Here, it is not straightforward to compute  $Z$  at all temperatures in a single exploration as the  $w_i$  and  $E_i$  terms in Eq. (10) depend on temperature. In the following sections, we specify the direct and the extended partition function for this case.

### 1. Direct method

The direct method, introduced by Szekeres et al<sup>21</sup>, consists in performing one exploration per temperature of interest, using the nested sampling algorithm. For each value of the inverse temperature  $\beta_j$ , the temperature-independent potential  $V(\mathbf{x}, \beta_j)$  is sampled according to the procedure presented in Section II B. The partition function is thus recovered in the same manner as Eq. (10), for each inverse temperature  $\beta_j$ :

$$Z_c(\beta_j) \approx \sum_i w_i e^{-\beta_j E_i(\beta_j)}, \tag{19}$$

where  $E_i(\beta_j)$  are the energies sampled in the exploration performed at fixed inverse temperature  $\beta_j$ . The internal energy and heat capacity are computed by discretizing Eqs. (17) and (18) using the fact that the average of a function  $g$  depending on  $\mathbf{x}$  and  $\beta$  can be computed as

$$\langle g(\mathbf{x}, \beta_j) \rangle \approx \frac{\sum_i w_i g(\mathbf{x}^{(i)}, \beta_j) e^{-\beta_j E_i(\beta_j)}}{Z_c(\beta_j)}, \tag{20}$$

where  $\mathbf{x}^{(i)}$  is the position vector at iteration  $i$ . The full expressions of  $U_c$  and  $C_{V,c}$  are given in the supplementary material (Section I A).

### 2. Extended partition function method

We now introduce a method allowing to obtain the same information for all temperatures in one exploration, as in the

case of temperature-independent potential. The key idea of the method is to distinguish between the temperature  $\beta$  at which the partition function is computed and the temperature which appears in the potential  $V$ , denoted hereafter by  $\tilde{\beta}$ , which we probe as an auxiliary parameter. Therefore, we sample both the positions  $\mathbf{x}$  and the auxiliary temperature  $\tilde{\beta}$ , instead of just the positions as in the direct method, to eventually converge to the same physical results.

We want to compute the partition function from Eq. (15). To that end, we first define the *extended partition function*  $\tilde{Z}_c(\beta)$  as

$$\tilde{Z}_c(\beta) = \int \int d\mathbf{x} d\tilde{\beta} e^{-\beta V(\mathbf{x}, \tilde{\beta})}. \tag{21}$$

In this formula, the temperature parameter  $\beta$  in  $Z_c(\beta)$  is split in two parameters which play distinct roles: the physical temperature  $\beta$ , at which the partition function is computed, and the auxiliary temperature  $\tilde{\beta}$ , which is a parameter used to sample the parameter space. The standard partition function  $Z_c(\beta)$  can be recovered from the extended partition function  $\tilde{Z}_c(\beta)$  via

$$Z_c(\beta) = \int \int d\mathbf{x} d\tilde{\beta} \delta(\beta - \tilde{\beta}) e^{-\beta V(\mathbf{x}, \tilde{\beta})} \tag{22}$$

$$= \langle \delta(\beta - \tilde{\beta}) \rangle_{\tilde{\beta}} \tilde{Z}_c(\beta), \tag{23}$$

where  $\langle \delta(\beta - \tilde{\beta}) \rangle_{\tilde{\beta}}$  is the average, according to the probability distribution  $\exp(-\beta V(\mathbf{x}, \tilde{\beta}))/\tilde{Z}_c(\beta)$ , of the Dirac delta-function  $\delta(\beta - \tilde{\beta})$  in the extended  $(\mathbf{x}, \tilde{\beta})$  space. In the above expression,  $Z_c(\beta)$  corresponds to the partition function at inverse temperature  $\beta$  of the potential  $V(\mathbf{x}, \tilde{\beta})$ , from which the delta-function only selects the cases where  $\beta = \tilde{\beta}$ . As in  $\tilde{Z}_c(\beta)$ , the potential  $V$  does not depend on the inverse temperature  $\beta$  but on the auxiliary inverse temperature  $\tilde{\beta}$ : it is therefore possible to compute  $Z_c(\beta)$  at all temperatures in a single exploration, as in the classical case, but at the cost of adding an extra parameter in the exploration, namely  $\tilde{\beta}$ . This method is thus referred to as the *extended partition function* method or *extended* method in short.

Ideally, to compute  $Z_c(\beta)$  at temperature  $\beta$ , we would like to use only the values of  $\tilde{\beta}$  for which  $\tilde{\beta} = \beta$ . However, in practice, because of the use of discrete sampling, a true delta-function is too restrictive: the probability to sample points for which  $\tilde{\beta} = \beta$  is null. We therefore use a function  $f(t; \alpha)$ , depending on a parameter  $\alpha$ , that approximates the delta function<sup>44</sup> in the  $\alpha \rightarrow \infty$  limit, i.e.,  $\lim_{\alpha \rightarrow \infty} f(t; \alpha) = \delta(t)$ . As a result, instead of selecting the values where  $\tilde{\beta} = \beta$ , we choose the points sampled at a temperature  $\tilde{\beta}$  where  $\tilde{\beta} \approx \beta$ : a weight is given to each value of  $\tilde{\beta}$  through the function  $f$  and used to compute  $Z_c(\beta)$ :

$$Z_c(\beta) = \langle \delta(\beta - \tilde{\beta}) \rangle_{\tilde{\beta}} \tilde{Z}_c(\beta) \approx \langle f(\beta - \tilde{\beta}; \alpha) \rangle_{\tilde{\beta}} \tilde{Z}_c(\beta). \tag{24}$$

In the following, the double-brackets  $\langle \langle \cdot \rangle \rangle_{\alpha}$  will denote

$$\langle \langle g(\mathbf{x}, \beta) \rangle \rangle_{\alpha} = \langle f(\beta - \tilde{\beta}; \alpha) g(\mathbf{x}, \tilde{\beta}) \rangle_{\tilde{\beta}} \tag{25}$$

the average of a function  $g$  according to the probability distribution  $\exp(-\beta V(\mathbf{x}, \tilde{\beta}))/\tilde{Z}_c(\beta)$  using  $f$  as the delta-function approximation with parameter  $\alpha$ . The average also depends on the choice of the prior on  $\tilde{\beta}$ , which will be discussed in more details in Section III B 1 for the path-integral formalism. Equation (25) can be discretized using

$$\langle\langle g(\mathbf{x}, \beta) \rangle\rangle_\alpha \approx \frac{\sum_i w_i g(\mathbf{x}^{(i)}, \tilde{\beta}_i) f(\beta - \tilde{\beta}_i; \alpha) e^{-\beta E_i}}{\tilde{Z}_c(\beta)} \quad (26)$$

with  $w_i$  as in Eq. (11) and  $\mathbf{x}^{(i)}$  the position vector,  $\tilde{\beta}_i$  and  $E_i$  the inverse temperature and energy sampled at iteration  $i$ . The term  $E_i = E(\mathbf{x}^{(i)}, \tilde{\beta}_i)$  is computed at inverse temperature  $\tilde{\beta}_i$  for the position vector  $\mathbf{x}^{(i)}$ . Note that, in practice, increasing the value of  $\alpha$ , while providing a better approximation of the delta function, results in a smaller window and therefore to small amounts of sampling points being used to compute the thermodynamic properties at each temperature. On the contrary, small values of  $\alpha$  use more sampling points but those points are sampled on a wider temperature range. The impact of the choice of  $\alpha$  and therefore of  $f(\beta - \tilde{\beta}; \alpha)$  in Eq. (25) will be discussed in Sections IV (harmonic potentials) and V (Lennard-Jones clusters).

Similarly, the internal energy and heat capacity are computed as (see Eqs. (17) and (18))

$$U_c(\beta) \approx \frac{\tilde{Z}_c(\beta)}{Z_c(\beta)} \left\langle \left\langle V(\mathbf{x}, \beta) + \beta \frac{\partial V(\mathbf{x}, \beta)}{\partial \beta} \right\rangle \right\rangle_\alpha \quad (27)$$

and

$$\begin{aligned} C_{v,c}(\beta) \approx & \frac{\tilde{Z}_c(\beta)}{k_B T^2 Z_c(\beta)} \left\langle \left\langle \left( V(\mathbf{x}, \beta) + \beta \frac{\partial V(\mathbf{x}, \beta)}{\partial \beta} \right)^2 \right\rangle \right\rangle_\alpha \\ & - \frac{\tilde{Z}_c(\beta)^2}{k_B T^2 Z_c(\beta)^2} \left( \left\langle \left\langle \left( V(\mathbf{x}, \beta) + \beta \frac{\partial V(\mathbf{x}, \beta)}{\partial \beta} \right) \right\rangle \right\rangle_\alpha \right)^2 \\ & - \frac{\tilde{Z}_c(\beta)}{k_B T^2 Z_c(\beta)} \left\langle \left\langle \left( 2 \frac{\partial V(\mathbf{x}, \beta)}{\partial \beta} + \beta \frac{\partial^2 V(\mathbf{x}, \beta)}{\partial \beta^2} \right) \right\rangle \right\rangle_\alpha, \end{aligned} \quad (28)$$

respectively. The full expressions of  $U_c$  and  $C_{v,c}$  are given in the supplementary material (Section I B).

From a computational point of view, with the extended method, the series of  $(w_i, E_i, \tilde{\beta}_i)$  is sampled once to compute the thermodynamic properties in a pre-determined temperature range, in contrast with the direct method where the sampling must be carried out separately for each temperature. While the sampling is more expensive for the extended method, we anticipate that the total number of required samples is smaller than for the direct method.

In this work, we consider two choices for the function  $f$  (both normalized to unity) among the many possible approximations of the Dirac delta-function:

- the rectangular function

$$f_r(\beta - \tilde{\beta}; \alpha) = \begin{cases} \alpha/2 & \text{if } |\beta - \tilde{\beta}| < 1/\alpha, \\ 0 & \text{otherwise;} \end{cases} \quad (29)$$

$f_r$  selects all sampled points for which  $\tilde{\beta}$  is  $1/\alpha$ -close to  $\beta$  giving them all the same weight.

- the Gaussian function

$$f_G(\beta - \tilde{\beta}; \alpha) = \frac{\alpha}{\sqrt{\pi}} e^{-\alpha^2(\beta - \tilde{\beta})^2}; \quad (30)$$

$f_G$  gives a normal weight to all sampled points: points for which  $\tilde{\beta}$  is distant from  $\beta$  of more than  $3\frac{1}{\sqrt{2}\alpha}$  (three standard deviations) will have a very small weight and participate little to  $Z_c^P(\beta)$ .

In both case,  $\alpha$  has the dimension of an energy.

### III. THE QUANTUM PARTITION FUNCTION IN THE PATH-INTEGRAL FORMALISM

#### A. General equations

In this section, we present the specific form of the partition function for the particular case of a temperature-dependent potential resulting from the treatment of NQEs via Feynman's path-integral formalism<sup>45</sup>. The quantum nuclei are represented by classical closed polymers composed of  $P$  replicas. When  $P \rightarrow \infty$ , the probability distribution for the replicas converges toward the exact quantum distribution. The partition function  $Z^P$  of this new system for  $N$  atoms in three dimensions and  $P$  replicas is<sup>21,46</sup>

$$Z^P(\beta) = \left( \frac{Pm}{2\pi\beta\hbar^2} \right)^{3NP/2} \int \cdots \int d\mathbf{x}_1 \dots d\mathbf{x}_P e^{-\beta V_P} \quad (31)$$

with the potential

$$V_P(\mathbf{x}_1, \dots, \mathbf{x}_P, \beta) = \underbrace{\frac{mP}{2\hbar^2\beta^2} \sum_{i=1}^P (\mathbf{x}_i - \mathbf{x}_{i+1})^2}_{Q(\mathbf{x}, \beta)} + \underbrace{\frac{1}{P} \sum_{i=1}^P V(\mathbf{x}_i)}_{\bar{V}(\mathbf{x})} \quad (32)$$

The term  $\bar{V}$  is the interaction potential  $V$  of the quantum nuclei averaged over all replicas and  $Q$  is a term that derives from the quantum kinetic operator and corresponds to an harmonic interaction between the replicas. We denote

$$\mathbf{x}_i = (x_{i,1}, y_{i,1}, z_{i,1}, \dots, x_{i,N}, y_{i,N}, z_{i,N}) \quad (33)$$

the position vector of the  $i$ -th replica of the system and  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_P)$  the full position vector. In that case, we have

$$\frac{\partial V_P(\mathbf{x}, \beta)}{\partial \beta} = -\frac{2}{\beta} Q(\mathbf{x}, \beta) \quad (34)$$

and

$$\frac{\partial^2 V_P(\mathbf{x}, \beta)}{\partial \beta^2} = \frac{6}{\beta^2} Q(\mathbf{x}, \beta). \quad (35)$$

When  $T \rightarrow \infty$ , the polymer collapses onto itself and resembles a classical particle. For this limiting case, a single replica

can then be used again. On the contrary, when  $T \rightarrow 0$ , the spring interaction between replicas decreases and the polymer spreads over: the number of replicas needed also increases in order better to represent the quantum positional uncertainty<sup>46</sup>.

By analogy with the classical case (Eq. (6)), we denote by  $Z_k^P$  the prefactor in Eq. (31):

$$Z_k^P(\beta) = \left( \frac{Pm}{2\pi\beta\hbar^2} \right)^{3NP/2}. \quad (36)$$

The corresponding contributions to the internal energy and to the heat capacity are<sup>21</sup>  $U_k^P(\beta) = \frac{3NP}{2\beta}$  and  $C_{v,k}^P(\beta) = \frac{3NPk_B}{2}$ . The position-dependent term is thus written

$$Z_c^P(\beta) = \int \cdots \int dx_1 \dots dx_P e^{-\beta V_p}, \quad (37)$$

where  $V_p$  is defined in Eq. (32). In general,  $Z_c$  does not have an analytical expression. We therefore use the nested sampling algorithm to compute it, either with the direct or with the extended method. Note that there are two distinct convergence parameters: the number of live points  $K$ , as in the classical case, and the number of replicas  $P$  that has to be large enough to correctly represent the NQEs.

In Sections II C 1 and II C 2, we have seen that, due to the temperature dependence of the potential, the discrete expressions to recover the internal energy and the heat capacity are more complex than those from Eqs. (12) and (13), as they include new terms appearing whenever the effective potential depends on  $\beta$ . In Appendix B, we explicitly provide those expressions, in both direct and extended methods, in the particular case of the path-integral potential  $V_p(x, \beta)$  (Eq. (32)). Implementing the direct method is straightforward: nested sampling explorations are performed at the temperatures of interest and the thermodynamic properties are computed for each exploration. However, the implementation of the extended method is more complex and is discussed in the next section.

## B. Extended path-integral algorithm implementation

### 1. Choice of variables to explore the extended space

Before using the extended method in practice, we present how to sample points using this method. Indeed, simply turning the auxiliary temperature  $\tilde{\beta} = 1/(k_B \tilde{T})$  into an additional parameter to explore in the nested sampling procedure with a uniform prior would result in an imbalanced sampling due to the particular form of the potential  $Q$  in Eq. (32). This would cause the low auxiliary temperatures  $\tilde{T}$  to be overrepresented in the sampling with respect to high  $\tilde{T}$ . This affects the efficiency of the method and would require extremely large numbers  $K$  of live points to converge. An example is provided in Figure 1 (a) for the harmonic potential. We see that the sampling of the temperature is imbalanced, with low temperatures being more sampled than high temperatures.

In order to overcome this problem, we change variable to obtain a more balanced sampling between low and high aux-

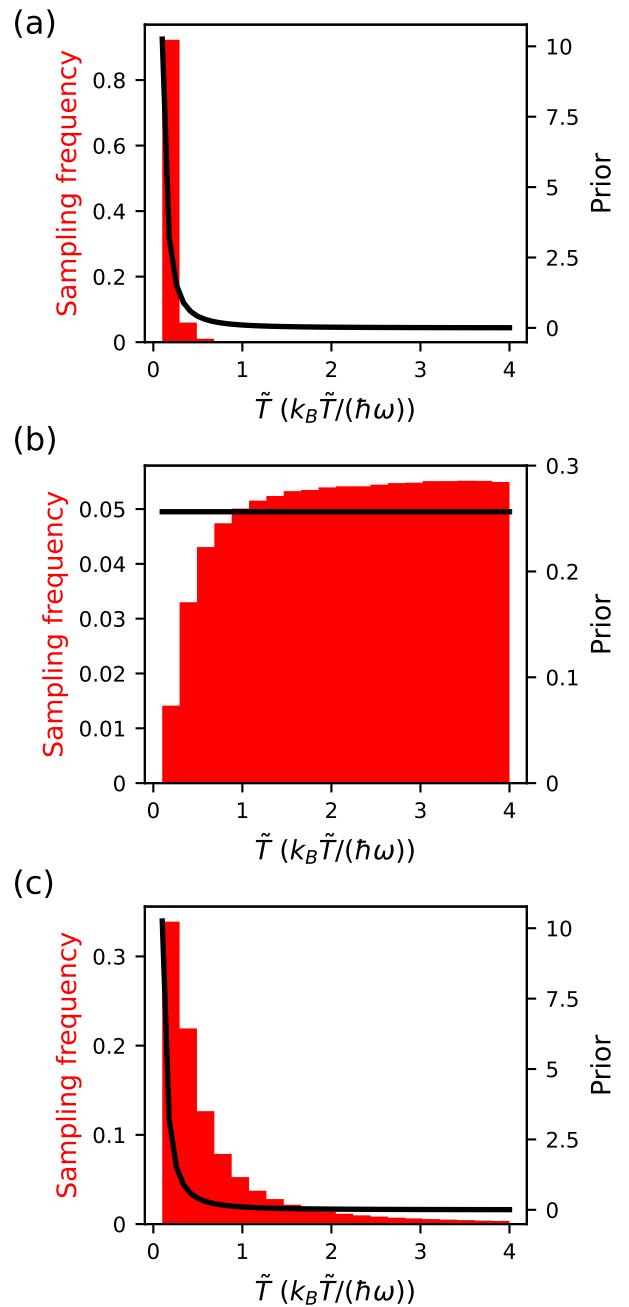


FIG. 1. Sampling frequency for the harmonic potential with  $N = 1$  and  $P = 2$  (a) before the change of variables using  $\tilde{\beta}$  as the explored parameter, (b) after the change of variables using  $1/\tilde{\beta}$  as the explored parameter and (c) after the change of variables using  $\tilde{\beta}$  as the explored parameter. We have  $\tilde{T} = 1/(k_B \tilde{\beta})$ . The prior is indicated by the black curve.

iliary temperatures. We denote

$$\tilde{y}_i = \frac{x_i - \bar{x}}{\lambda_P(\tilde{\beta})} \quad 1 \leq i \leq P, \quad (38)$$

with  $\bar{x} = \frac{1}{P} \sum_{i=1}^P \mathbf{x}_i$  the centroid of the replicas and

$$\lambda_P^2(\tilde{\beta}) = \frac{\hbar^2 \tilde{\beta}^2}{mP}. \quad (39)$$

The variable  $\tilde{y}_i$  is the position of the  $i$ -th replica with respect to the centroid, normalized by the factor  $\lambda_P$ , used to remove the temperature dependence of the harmonic interaction between replicas, as can be seen in the resulting expression for the potential

$$V_P(\mathbf{x}_1, \dots, \mathbf{x}_P, \tilde{\beta}) = V_P(\bar{x}, \tilde{y}_1, \dots, \tilde{y}_P, \tilde{\beta}) \quad (40)$$

$$= \frac{1}{2} \sum_{i=1}^P (\tilde{y}_i - \tilde{y}_{i+1})^2 + \frac{1}{P} \sum_{i=1}^P V(\bar{x} + \lambda_P(\tilde{\beta}) \tilde{y}_i). \quad (41)$$

In this case, the position parameters are the position of the centroid  $\bar{x}$  and of the first  $P-1$  replicas  $\{\tilde{y}_i\}_{1 \leq i < P}$ . The  $P$ -th replica can be determined from the others as  $\tilde{y}_P = -\sum_{i=1}^{P-1} \tilde{y}_i$ . The auxiliary temperature dependence has been moved from the interaction between replicas  $Q$  to the potential energy term  $V$ . Figure 1 (b)–(c) shows the sampling distribution after the transformation using  $\tilde{\beta}$  and  $1/\tilde{\beta}$  as the extra parameters with uniform prior. We see that the change of variable in the prior distribution results in a more balanced sampling. Furthermore, the sampling frequency approaches the prior very closely when sampling with the uniform prior on  $1/\tilde{\beta}$  (Figure 1 (c)). This shows that the choice of the function of  $\tilde{\beta}$  to explore with uniform prior is also important. It is important to note that this change of variables and the choice of the parameter with uniform prior does not affect the fully converged results that would be obtained in the limit of an infinite number of live points  $K$ . However, an appropriate choice of variables can greatly reduce the number of live points needed in practice to converge and therefore the computational cost of the exploration.

## 2. Choice of the upper and lower bounds of the parameters

With the change of variables presented above, the temperature  $\tilde{\beta}$  affects the distance between the replicas and the centroid. `Nested_fit` requires that we fix a pre-determined interval in which the parameters evolve: for high  $\tilde{\beta}$  (low temperatures), the replicas can go further from the centroid than for low  $\tilde{\beta}$  (high temperatures). Hence, choosing the lower and upper bounds of the parameters is not as straightforward as for the direct method. Indeed, without transformation, we simply take all replicas to be in an identical box of fixed size. With the transformation, the centroid is also placed in a box of fixed size. The difficulty arises when choosing the bounds for the  $\tilde{y}_i$ , that is, the position of the replicas relative to the centroid. Indeed, if  $\tilde{y}_i$  is in a cubic box of side  $a$  centered in 0, then the corresponding replica  $\mathbf{x}_i$  is in a cubic box of size  $a\lambda_P(\tilde{\beta})$  centered in  $\bar{x}$ : the volume of space that is sampled by the replicas therefore depends on the temperature. The impact of the parameter transformation on the selection of the box size is represented in Figure 2. We will indicate how to choose these

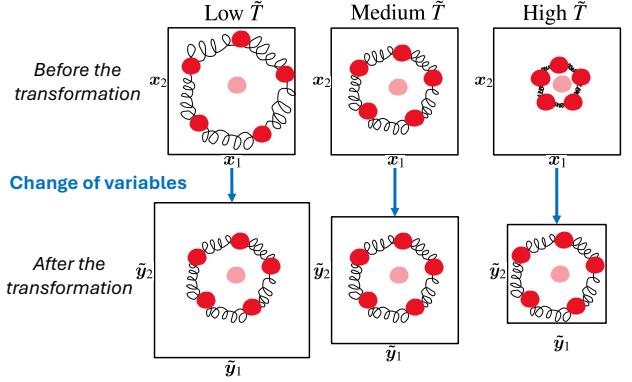


FIG. 2. Representation of the parameter transformation given by Eqs 39–41 (for  $P = 5$ ). Before the transformation, the polymer is in a box of fixed size (the same for all temperatures). In that case, the extension of the polymer depends on the temperature: the polymer contracts with increasing temperature. After the transformation, the extension of the polymer is the same for all temperatures (see first term of Eq. 41), however the size of the box is then proportional to  $\lambda_P$ , hence to the temperature.

bounds in practice for the two examples studied in this work: the quantum harmonic potential and Lennard-Jones clusters.

## IV. HARMONIC SYSTEMS

For the quantum harmonic potential, the exact expressions for the internal energy and the heat capacity are known analytically<sup>46,47</sup> (see Section IV A), which makes it a good test for the implementation of nested sampling for studying NQEs.

### A. Exact equations

As a first application of the path-integral nested sampling method, we consider the following harmonic potential for one particle ( $N = 1$ ):

$$V(\mathbf{x}) = \frac{m\omega^2}{2} (x^2 + y^2 + z^2), \quad (42)$$

with  $\mathbf{x} = (x, y, z)$  the position vector of the particle. Consequently, the expression of  $V_P$  in Eq. (32) gives

$$V_P(\mathbf{x}_1, \dots, \mathbf{x}_P, \beta) = \frac{mP}{2\hbar^2 \beta^2} \sum_{i=1}^P (\mathbf{x}_i - \mathbf{x}_{i+1})^2 + \frac{1}{P} \sum_{i=1}^P \frac{m\omega^2}{2} (x_i^2 + y_i^2 + z_i^2), \quad (43)$$

where  $m$  is the particle mass,  $\omega$  the frequency and  $\mathbf{x}_i = (x_i, y_i, z_i)$  is the three-dimensional position vector of the  $i$ -th replica. In practice, we use reduced units, i.e., we take  $m$  and  $\omega$  to be the units of mass and frequency. Hence, in reduced units, we have  $m = 1$  and  $\omega = 1$ . We also take  $\hbar = 1$ . Furthermore, we use the reduced temperature  $\theta = k_B T / (\hbar\omega)$  and

reduced dimensionless length units by normalizing positions by the factor  $\sqrt{\hbar/(m\omega)}$ . For the harmonic potential, the values of the lower and upper bounds of the parameters are not important as long as the sampling space is big enough not to be affected by finite size effects at the temperatures we are sampling (see Ref. 39 for the study in the classical case). This is not the case for all potentials, as we will see when studying Lennard-Jones clusters: indeed, the harmonic potential is a confining potential, contrary to the Lennard-Jones potential which is null at infinite interparticle distance. Therefore, for the direct method, all replicas are placed in a finite box of size  $L = 30$  in reduced units; for the extended method, the centroid and the  $\tilde{y}_i$  ( $1 \leq i \leq P-1$ ) are also confined in a box of side  $L = 30$  in reduced units.

In one dimension, the system can take the energy values<sup>46</sup>  $E_n = (n + 1/2)$  for  $n = 0, 1, 2, \dots$ , corresponding to the eigenvalues of the Hamiltonian  $H$ . In the quantum setting, the partition function is given by<sup>46</sup>  $Z(\beta) = \text{Tr}(\exp(-\beta H))$ , hence

$$Z(\beta) = \sum_{n=0}^{\infty} e^{-\beta E_n} = \frac{1}{2 \sinh\left(\frac{1}{2\theta}\right)}. \quad (44)$$

Consequently, we can compute the exact internal energy and heat capacity: for a three-dimensional system with one particle, the internal energy and heat capacity of the one-dimensional particle are multiplied by 3, giving<sup>46,47</sup>

$$U(\beta) = \frac{3}{2} \coth\left(\frac{1}{2\theta}\right) \quad (45)$$

and

$$C_v(\beta) = \frac{3}{4\theta^2 \sinh^2\left(\frac{1}{2\theta}\right)} k_B, \quad (46)$$

respectively.

Furthermore, for the harmonic potential, we can compute analytically the exact values of the heat capacity for a finite value of  $P$  in particular,

$$C_v^{P=2}(\beta) = 3k_B \left( 2 - \frac{4\theta^4 + \frac{3}{4}\theta^2}{(2\theta^2 + \frac{1}{8})^2} \right) \quad (47)$$

and

$$C_v^{P=4}(\beta) = 3k_B \left( 4 - 2 \frac{\theta^4 + \frac{3}{32}\theta^2}{(\theta^2 + \frac{1}{32})^2} - \frac{4\theta^4 + \frac{3}{16}\theta^2}{(2\theta^2 + \frac{1}{32})^2} \right) \quad (48)$$

for  $P = 2$  and  $P = 4$ , respectively. The derivation of these expressions is given in Appendix C.

In contrast with the classical case, the quantum heat capacity in Eq. (46) goes to zero in the low temperature limit. However, for finite numbers of replicas, the expressions (47) and (48) still tend to finite value for  $T \rightarrow 0$ , with  $C_v^{P=2} \rightarrow 6k_B$  for  $P = 2$  and  $C_v^{P=4} \rightarrow 12k_B$  for  $P = 4$  for one particle in three dimensions. This corresponds to the classical heat capacity for a system with  $6P$  degrees of freedom.

Before using `nested_fit` with both the direct and extended methods on the harmonic system, we first provide a theoretical test of the effect of the use of a smeared delta-function  $f(\beta - \tilde{\beta}; \alpha)$  in the extended method.

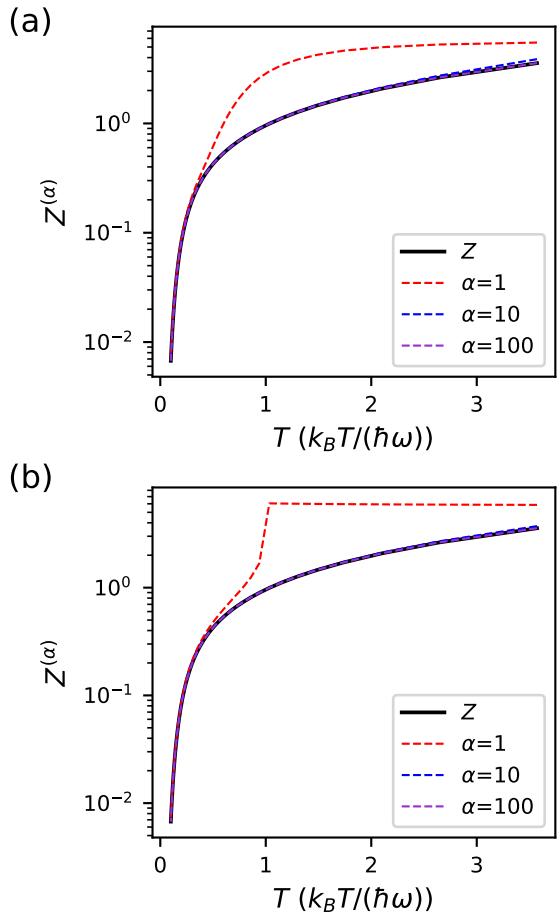


FIG. 3. Comparison between  $Z$  and the smeared-delta function approximation for (a) the Gaussian function (Eq. (50)) and (b) the rectangular function (Eq. (51)). The curves corresponding to  $Z$ ,  $\alpha = 10$  and  $\alpha = 100$  almost coincide at the graph scale.

## B. Theoretical test of smearing functions

In this section, we compute analytical results on the effect of the two smeared delta-functions  $f(\beta - \tilde{\beta}; \alpha)$  previously introduced. For that purpose, we use the quantum harmonic potential: for simplicity, we consider the one-dimensional case instead of the three-dimensional one. The exact partition function<sup>46</sup> is given in Eq. (44) and can be rewritten as an integral over  $\tilde{\beta}$  as in Section II C 2:

$$Z(\beta) = \int_0^\infty d\tilde{\beta} \delta(\beta - \tilde{\beta}) Z(\tilde{\beta}). \quad (49)$$

Replacing the delta-function by the Gaussian window gives the approximation

$$Z^{(\alpha)}(\beta) = \int_0^\infty d\tilde{\beta} \sqrt{\frac{\alpha}{\pi}} e^{-\alpha(\beta - \tilde{\beta})^2} Z(\tilde{\beta}), \quad (50)$$

which depends on parameter  $\alpha$ . Replacing it by the rectangular window gives

$$Z^{(\alpha)}(\beta) = \int_{\max(0, \beta - 1/\alpha)}^{\beta + 1/\alpha} d\tilde{\beta} \frac{\alpha}{2} Z(\tilde{\beta}). \quad (51)$$

In the case of Eq. (50), we compute  $Z$  numerically. In the case of the rectangular window (Eq. (51)), the integral yields

$$Z^{(\alpha)}(\beta) = \frac{\alpha}{2\hbar\omega} \left[ \log \left( \tanh \left( \frac{\hbar\omega(\beta + 1/\alpha)}{4} \right) \right) - \log \left( \tanh \left( \frac{\hbar\omega \max(0, \beta - 1/\alpha)}{4} \right) \right) \right]. \quad (52)$$

In order to compare the exact and the extended method results for the Gaussian window, we represent the curves for different values of  $\alpha$  in Figure 3 (a). We take values of  $k_B T = 1/\beta$  between  $0.2\hbar\omega$  ( $\beta = 5$ ) and  $3.5\hbar\omega$  ( $\beta = 0.28$ ) which is the range of temperature studied in Section IV C. In Figure 3, we see that if we take  $\alpha$  large enough ( $\alpha \gtrsim 10$ ),  $Z$  is correctly recovered for the range of temperature studied. For smaller values of  $\alpha$ , the curve is not correctly recovered for high temperatures. This is likely due to the fact that, for these combinations of  $\alpha$  and  $\beta$ , the delta-function approximation gives significant weight to small values of  $\tilde{\beta}$  where the integrand diverges. However, at high temperature a classical treatment of the system is possible (in that case at around  $\sim 2.5k_B T/(\hbar\omega)$ , see Figure 4), so that the behavior of  $Z$  can be obtained by the simpler, classical version of the algorithm.

In Figure 3 (b), we represent the curves for different values of  $\alpha$  for the same system, using the rectangular function. In that case, since  $\log(\tanh(x))$  is not defined in 0, we replace  $\max(0, \beta - 1/\alpha)$  by  $\max(10^{-5}, \beta - 1/\alpha)$  in Eq. (52) to avoid numerical problems. This approximation can be done as, in practice, we sample between finite values of  $\tilde{T}$  ( $\tilde{\beta} > 0$ ). We observe a similar behavior with increasing  $\alpha$  than what was observed for the Gaussian window. In that case too, we have a good recovery of  $Z$  for  $\alpha \gtrsim 10$ .

### C. Direct method - comparison of `nested_fit` with the exact cases

Figure 4 presents the results obtained for the heat capacity for  $P = 1, 2, 4, 8$  with `nested_fit` using the direct method (the exploration parameters are given in Appendix D, as will be the case for all applications in this work). The exact quantum, classical ( $P = 1$ ),  $P = 2$  and  $P = 4$  curves are also represented. As check, we perform one exploration per temperature for the  $P = 1$  curve with `nested_fit`, even though this case is equivalent to the classical case and only one exploration can be performed for all temperatures. We can see that nested sampling is able to recover the exact curves. We also see that when  $P$  increases, the exact quantum curve is recovered down to lower temperatures, which is the expected behavior as the system is, so to say, "more quantum" at lower temperatures and thus requires more replicas to converge. Moreover, we see that at low temperatures, the curves for  $P = 2$  and  $P = 4$  (both exact and obtained via `nested_fit`) start increasing, with the minimum of the curve being reached at a lower temperature for  $P = 4$  than for  $P = 2$ . The unphysical increase in the heat capacity is the manifestation of the classical nature of the system for a finite number of replicas, whereas the limit  $P \rightarrow \infty$  would be needed to simulate the quantum systems

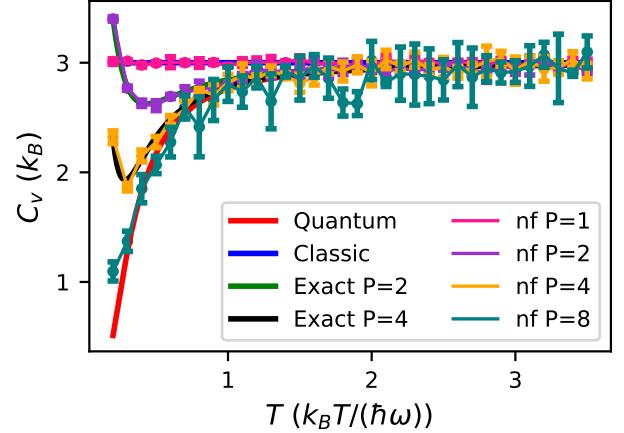


FIG. 4. Heat capacity for the quantum harmonic potential with the direct method. The analytical results for the classical and quantum cases are plotted as full lines, as well as the formulas for two (Eq. (47)) and four (Eq. (48)) replicas ( $P = 2$  and  $P = 4$ , respectively). Also shown are the numerical results obtained with `nested_fit` (nf in the legend). The points correspond to the temperatures at which the heat capacities were computed. Taking  $v = \omega/(2\pi) = 100\text{cm}^{-1}$ , which is a typical vibration frequency in solids, the temperature range considered here would correspond to [290K,5000K]. The error bars were computed from four runs: for  $P = 1, 2, 4$ , they are smaller than the symbols for most temperatures.

down to  $T = 0$ . Therefore, for a finite number of replicas  $P$ , there is a threshold temperature  $T_P^*$  such that  $\partial C_v / \partial T < 0$  for  $T < T_P^*$ . Note that  $T_P^* \rightarrow 0$  as  $P \rightarrow \infty$ . Finally, we observe that the nested sampling results fluctuate more and have a bigger statistical uncertainty for higher  $P$ , which is again expected as a higher value of  $P$  means more degrees of freedom to explore. To reduce fluctuations, a larger number of live points  $K$  is required (here,  $K = 1000$  for all  $P$ ). The impact of  $K$  on the recovery of the heat capacity will be studied in Section V for the more complex partition function of Lennard-Jones clusters, which has multiple minima.

### D. Extended partition function method

#### 1. Choosing the smeared delta function

We now perform simulations with `nested_fit` using the extended method. For that, we use the path-integral harmonic potential with  $P = 2$ . The computational details are given in Appendix D.

For both approximations of the delta-function, we test different values of the parameter  $\alpha$ , corresponding to the size of the window where the inverse temperature  $\tilde{\beta}$  is sampled. In order to quantify the computational accuracy, we adopt the Mean Absolute Percentage Error (MAPE)<sup>48</sup>:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n 100 \frac{|S_i - E_i|}{E_i}, \quad (53)$$

with  $n$  the number of temperatures sampled,  $E_i$  the expected

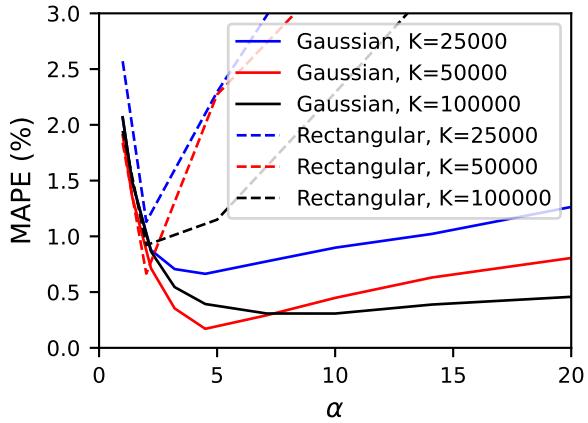


FIG. 5. MAPE as a function of  $\alpha$  for different values of  $K$  with the Gaussian (full lines) and rectangular (dashed lines) windows.

results — here the exact  $P = 2$   $C_V$  curve — and  $S_i$  the sampled points — here the computed  $C_V$  curve. In Figure 5, we present the MAPE curves as a function of  $\alpha$  for different values of  $K$  for both Gaussian and rectangular windows. The  $\alpha$  value for which the MAPE is the smallest changes with the window and number of live points  $K$ : for the Gaussian window, the smallest value of MAPE is obtained for  $\alpha = 4.5$  for  $K = 25000, 50000$  and  $\alpha = 10$  for  $K = 100000$  while for the rectangular window, the smallest MAPE value is obtained for  $\alpha \approx 2$  for all  $K$ . In the following, we refer to this value of  $\alpha$  as the "optimal" value, which is valid for a specific example and number of live points  $K$ . We point out that nested sampling provides a set of sampled points. These points are then post-processed by using distinct  $g(\beta - \tilde{\beta}; \alpha)$  windows: they are therefore issued from the same statistical set.

In Figure 6, we present the results obtained for the Gaussian and rectangular windows for  $\alpha = 3.2, 4.5, 7.1$  and  $\alpha = 1, 2, 5$  respectively, at  $K = 50000$ , for which the smallest MAPE value is obtained. Again, all curves were obtained from the same exploration. In both cases, taking a small value (below their optimal one) of  $\alpha$  does not correctly recover the heat capacity. Indeed, small values of  $\alpha$  correspond to large temperature windows, which gives rise to a bias. For large  $\alpha$  instead, we only select points that have an inverse temperature  $\tilde{\beta}$  very close to  $\beta$ . Since the number of selected points decreases when  $\alpha$  increases, large values of  $\alpha$  result in a poor statistical sampling and large fluctuations (without bias). We see that the rectangular function fluctuates more than the Gaussian function. A possible reason for this behavior is that it either selects or rejects the sampled points, with the selected points all having equal weights. The rectangular window function therefore does not discriminate between different inverse temperatures  $\tilde{\beta}$  depending on their distance from the target  $\beta$ , which is the case of the Gaussian function. Hence, from now on, we choose the Gaussian window for comparing the extended and direct methods.

In Figure 7, we represent the auxiliary temperature against the energy of the sampled configurations, colored according

to the weight given by the Gaussian approximation. We see that only a small proportion of points contribute significantly to the computation of the thermodynamic properties at a given temperature: this explains why we need such a huge number of live points to compute the thermodynamic properties at all temperatures with the extended method. Note that we have only represented the points for  $\tilde{T} \leq 1.5(k_B \tilde{T})/(\hbar\omega)$  but we have sampled up to  $\tilde{T} = 4(k_B \tilde{T})/(\hbar\omega)$ .

## 2. Comparing with the direct method

We now look at the performance of the extended method on the harmonic potential and compare our results with that obtained with the direct method (Figure 4). We again look at the curves obtained with `nested_fit` for  $P = 2, 4, 8$  as well as for  $P = 16$ . We put a uniform prior on  $\tilde{T}$  for  $P = 2, 8, 16$  and a uniform prior on  $\tilde{\beta}$  for  $P = 4$ . For the latter, the minimum of the  $P = 4$  exact curve (around  $0.3k_B T/(\hbar\omega)$ ) was not correctly recovered when using a uniform prior on  $\tilde{T}$ . This shows that the choice of the prior, which, contrary to  $\alpha$ , has to be chosen before the exploration, has a noticeable impact on the resulting curves as was illustrated in Section III B 1. The results are presented in Figure 8. We see that, with the extended method, `nested_fit` is able to recover the exact curves very well for  $P = 2, 4$ ; for  $P = 8, 16$  (for which we do not have reference curves), nested sampling converges to the exact quantum results (as expected for  $P \rightarrow \infty$ ). The corresponding sampling frequency are given in the supplementary material (Figure 1).

Finally, we compare the computational performance of the extended and direct methods on the  $P = 2$  case. We present the results in Figure 9 using  $K = 25000$  live points for the direct method (one exploration per temperature) and  $K = 50000$  for the extended one (one exploration for all temperatures) with  $\alpha = 7.1$ . Those choices result in similar MAPE values: 0.30% for the direct method and 0.29% for the extended method. Both methods correctly recover the exact curve and have error of the same order at high temperatures. At lower temperatures ( $T \leq 0.6(k_B T)/(\hbar\omega)$ ), the extended method has larger error than the direct method by around a factor 10: we have less samples at those temperatures (see supplementary material, Figure 1 (a)). To obtain the heat capacity curves for 34 temperatures ranging from  $k_B T = 0.2 \frac{k_B T}{\hbar\omega}$  to  $k_B T = 3.5 \frac{k_B T}{\hbar\omega}$ , the direct method requires  $2.9 \times 10^9$  energy evaluations (i.e., approximately  $8 \times 10^7$  energy evaluations per exploration) and the extended method requires  $1.8 \times 10^8$  energy evaluations. The extended method is thus approximately eight times faster than the direct method. We observe that, even though we use more live points, the extended method requires much less evaluations — by around a factor ten — than the direct one. Indeed, in the former case, we only need a single exploration to compute the heat capacity at all temperatures while we need multiple explorations in the latter case. Furthermore, if we want to add a point to the curve within the studied temperature range, a new exploration must be done for the latter while for the former the heat capacity for the new temperature can be computed from the previously sampled points as the choice of a

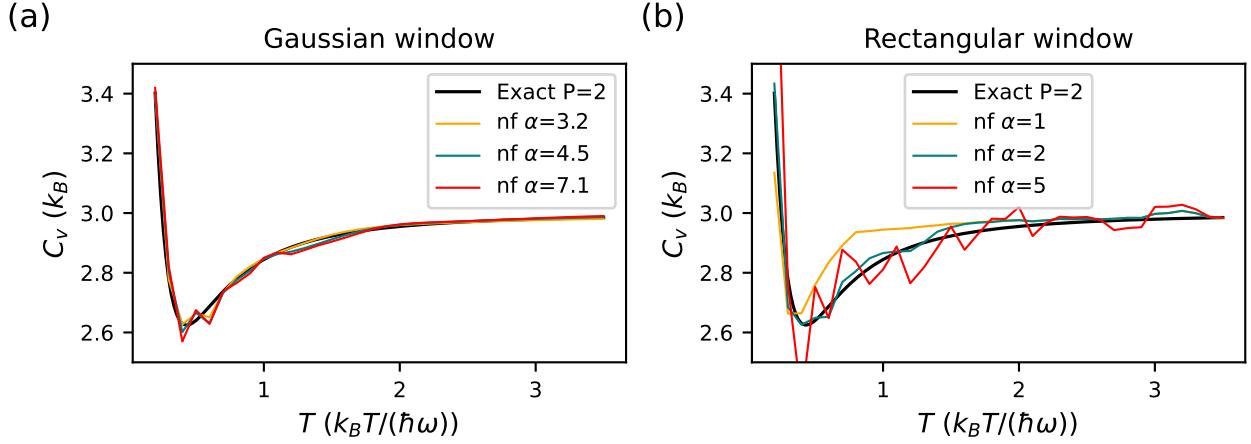


FIG. 6. Comparison of the heat capacity obtained for the harmonic potential with the extended method for different values of  $\alpha$  (a) for the Gaussian window and (b) the rectangular window. In each case only three values are represented for the readability of the graph.

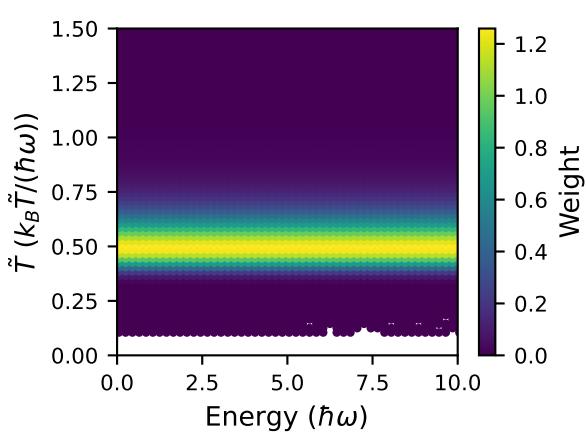


FIG. 7. Plot of the auxiliary temperature  $\tilde{T}$  against the energy of the sampled configurations for the  $P = 2$  and  $K = 50000$  case. The points are colored according to the weight given by the Gaussian approximation with  $\alpha = 5$  and  $T = 0.5(k_B T)/(\hbar\omega)$ .

specific temperature window is a post-processing step. Note that the two energy functions are slightly different as, in the extended case, there is a change of variables which is not used with the direct method. This change of variables results in a small increase in computational cost for the function, that becomes negligible when the number of parameters increases or the potential is more complex than the harmonic potential.

In conclusion, in this section, we have tested the extended partition function method on the harmonic potential. Mainly, we have looked at the impact of the choice of the smeared delta-function and of the corresponding parameter  $\alpha$ : we have seen that the Gaussian window leads to reduced fluctuations with respect to the rectangular one and that the optimal value of  $\alpha$  depends on the number of live points used and the system studied. Furthermore, comparing with the direct method, we have seen that the extended method requires less energy eval-

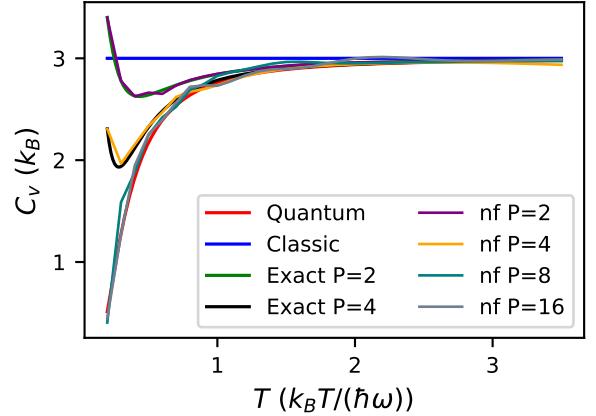


FIG. 8. Heat capacity for the quantum harmonic potential obtained using `nested_fit` (nf) with the extended method with  $\alpha = 4.5$  for  $P = 2$  and  $\alpha = 7.1$  for  $P = 4, 8, 16$ . We used  $K = 50000$ ,  $K = 100000$ ,  $K = 100000$  and  $K = 200000$  for  $P = 2$ ,  $P = 4$ ,  $P = 8$  and  $P = 16$ , respectively.

uations, roughly by an order of magnitude. Finally, we have seen that the choice of the window and  $\alpha$  is a post-processing step, contrary to the choice of the prior and number of live points  $K$ . This means that different  $\alpha$  can be tested on one exploration but that a new exploration must be performed to test a new prior (more details on the choice of the prior will be given in Section V D). In the next section, we study the impact of  $K$  on the recovery of the curves with both the direct and extended methods on the Lennard-Jones clusters that display a very large number of local energy minima.

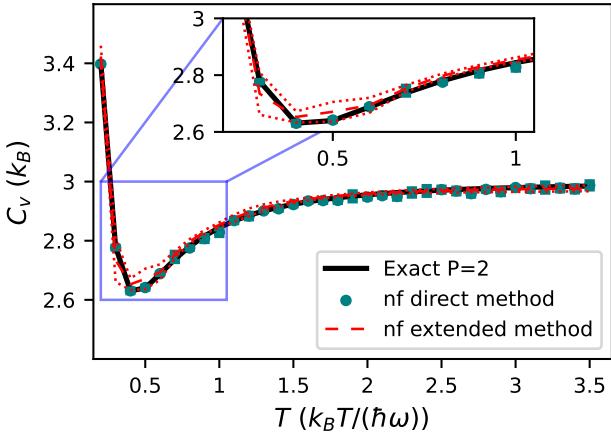


FIG. 9. Comparison of the direct and extended methods with  $\alpha = 7.1$  on the quantum harmonic potential with  $P = 2$ . The points correspond to the temperatures at which the heat capacities were computed with the direct method. The dotted red lines correspond to the first standard deviation for the extended run. The statistical uncertainties were computed from four runs.

## V. LENNARD-JONES CLUSTERS

### A. The potential model

Lennard-Jones clusters are formed by  $N$  particles interacting via

$$V(\mathbf{x}_i) = \sum_{\substack{1 \leq k < j \leq N \\ r_{i,kj} < r_c}} [V_{\text{LJ}}(r_{i,kj}) - V_{\text{LJ}}(r_c)] \quad (54)$$

with  $r_c$  the cutoff radius, which removes interaction at infinite range,  $r_{i,kj}$  the distance between the  $i$ -th replica of the  $k$ -th and  $j$ -th atoms and

$$V_{\text{LJ}}(r) = 4\epsilon \left( \left( \frac{r_0}{r} \right)^{12} - \left( \frac{r_0}{r} \right)^6 \right). \quad (55)$$

We have that  $r_0$  and  $\epsilon$  are the parameters of the potential that depend on the considered atomic species. We take  $r_c = 3r_0$ . Furthermore, we use periodic boundary conditions in a large cubic simulation box of size  $L \gg r_0$ .

We adopt reduced units: the temperature is in units of  $\frac{k_B T}{\epsilon}$  with  $\epsilon = 1$ . To fix the size of the box, we consider the value of the density  $\tilde{\rho}$  in unit of  $r_0^{-3}$ :  $\tilde{\rho} = \frac{N}{(L/r_0)^3} r_0^{-3}$ . We take  $L = 6$  and use  $r_0$  to tune the density to a given desired value. The density thus characterizes the size of the space the particles evolve in, relative to the number of atoms  $N$ , which is fixed. In this work,  $L$  is chosen so that the minimum image convention applies.

In the classical case, all Lennard-Jones systems are equivalent up to a rescaling of energies and distances, thus the results in reduced units are valid for all atomic species<sup>43</sup> by inserting the actual value for  $\epsilon$  and  $r_0$ . This is no longer true for the quantum case as changing the atomic species (and in particular the nuclei mass) changes the importance of the NQEs.

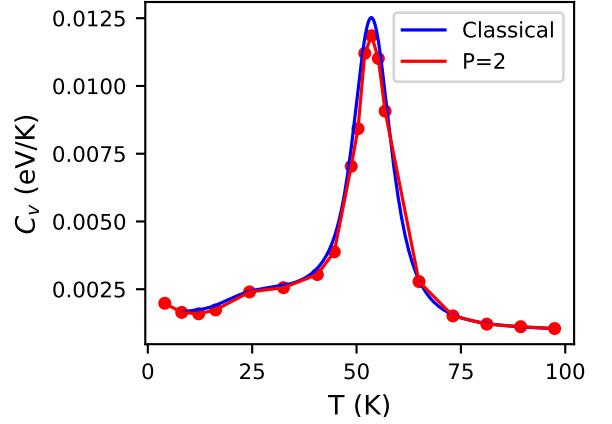


FIG. 10. Heat capacity for the 7-atom Krypton cluster with the direct method. The points correspond to the temperatures at which the  $P = 2$  heat capacity was computed.

This can be captured by the de Boer parameter<sup>49,50</sup>:

$$\Lambda = \frac{\lambda_B}{r_0}. \quad (56)$$

The parameter  $\Lambda$  is the ratio between an effective the Broglie wavelength<sup>49,51</sup>  $\lambda_B = \frac{\hbar}{\sqrt{m\epsilon}}$ , for particles of mass  $m$  and typical energy  $\epsilon$ , and a characteristic length  $r_0$  (of the order of the interparticle distance). Hence, a high value of parameter  $\Lambda$  ( $\Lambda \gtrsim 1$ ) indicates that the quantum effects are important while a small value ( $\Lambda \ll 1$ ) indicates that they are less important and that a classical treatment of the system may be suitable. As we can see from Eq. (56), the de Boer parameter depends on the mass. Therefore, in the following, for each example, we will fix  $\hbar = 1$  ( $h = 2\pi$ ) as well as the values of  $\epsilon$  and  $r_0$  and change the mass to have the de Boer parameter of the studied atomic species.

### B. 7-atom Krypton cluster Kr<sub>7</sub>

We first consider a 7-atom Krypton cluster, which has a de Boer parameter of<sup>52</sup>  $\Lambda = 0.10$ . In that case, we take  $\epsilon = 1$ ,  $r_0 = 0.659$  — this corresponds to the classical case studied in Ref. 43 — and a box of size  $L = 6$  ( $\approx 9.1r_0$ ) for all replicas. From Eq. (56), this fixes  $m = \hbar^2/(\Lambda^2 r_0^2 \epsilon) \approx 9091$  in reduced units. In Figure 10, we represent the classical and path-integral heat capacities as obtained with two replicas in the direct method. We see that the two curves almost coincide, within the statistical error. Hence, in this case, the classical treatment of the atoms is possible as using two replicas gives sensibly the same results as using one. This behavior is in accordance with the small de Boer parameter of Krypton.

The main peak in the graph at  $T \sim 50\text{K}$  corresponds to the vapor-liquid transition, while the shoulder at  $T \sim 25\text{K}$  can be associated to a liquid-like to solid-like state<sup>17</sup>. The fact that  $C_v$  starts to increase again at low  $T$  for the path-integral curve is related to the finite value of  $P$ , as noted ahead in the harmonic oscillator case.

### C. 3-atom Neon cluster $\text{Ne}_3$

We then consider the case of a 3-atom Neon cluster, for which the de Boer parameter is<sup>52</sup>  $\Lambda = 0.59$ . We take  $\varepsilon = 1$ ,  $r_0 = 1$  and a box of size  $L = 6r_0$  for all replicas. Since  $\Lambda = 0.59$ , we take  $m = 113$ . This system is used to study the convergence of the heat capacity with the number of replicas and with the number of live points, first with the direct method and then with the extended method.

#### 1. Direct method

First, for each  $P$ , we study the convergence with the number of live points  $K$ , starting by<sup>53</sup>  $K > 3NP$  ( $3NP$  is the number of degrees of freedom). We then double this value until we obtain two consecutive heat capacity curves that look similar, that is, between which the MAPE (using the explorations with higher  $K$  as reference) is less than 1.25% for  $P = 2$ , 1.5% for  $P = 4$  and 3% for  $P = 8$ . The highest  $K$  value considered is  $K = 8000$  for  $P = 2$ ,  $K = 16000$  for  $P = 4$  and  $K = 32000$  for  $P = 8$ . Because for this system the exact analytical result is unknown, we consider the curve with the most live points and compare it to the curves obtained for smaller  $K$ . Again, we use the MAPE, with the curve obtained with the highest  $K$  as reference. We set a threshold of 5% for considering a curve to be converged. The MAPE is represented in Figure 11 as a function of  $K$  and for each value of  $P$ .

We can see that, for the same value of  $K$ , the error is higher for higher  $P$ , as expected, as higher  $P$  results in a higher number of degrees of freedom. In each case, we take as the first converged results the first point that is below the 5% threshold (dashed line):  $K = 64$  for  $P = 2$ ,  $K = 1024$  for  $P = 4$  and  $K = 16000$  for  $P = 8$ . We denote by  $K_5$  the number of live points needed for reducing the MAPE below the 5% threshold for  $P$  replicas. The curves obtained for various  $K$  values

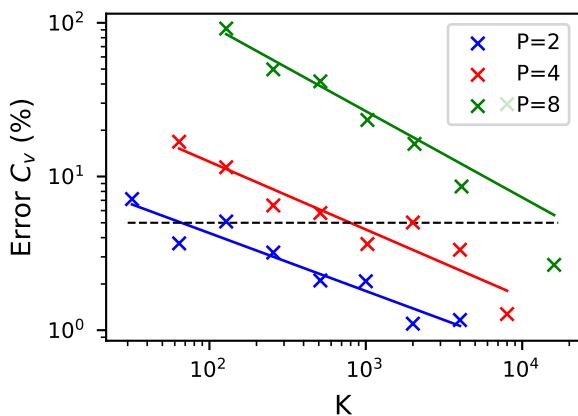


FIG. 11. MAPE error for the 3-atom Neon cluster with  $P = 2, 4, 8$  with the direct method. The black dashed line represents an error of 5%. The full lines are the linear log-log fits for each value of  $P$ . The slopes are  $-0.377$ ,  $-0.441$  and  $-0.561$  for  $P = 2$ ,  $P = 4$  and  $P = 8$ , respectively.

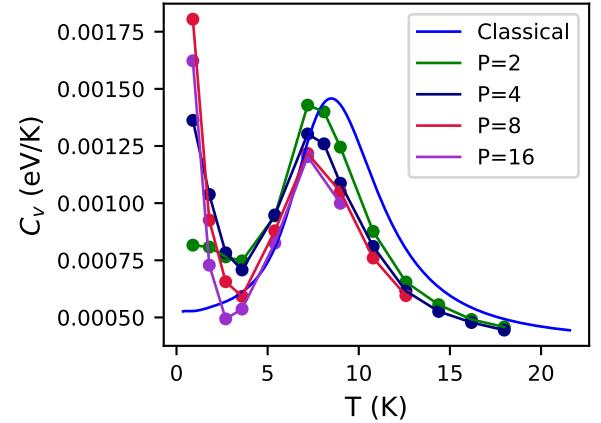


FIG. 12. Heat capacity for the 3-atoms Neon cluster. The explorations with  $K = 8000$ ,  $K = 16000$ ,  $K = 32000$  and  $K = 64000$  are represented for  $P = 2$ ,  $P = 4$ ,  $P = 8$  and  $P = 16$ , respectively.

for  $P = 2, 4, 8$  are given in the supplementary material (Figure 2). Finally, we expect the statistical uncertainty to scale as  $1/\sqrt(K)$ <sup>15,54</sup> and we indeed see that the MAPE decreases with the value of  $K$  with a slope of  $\sim -0.5$  in log-log scale.

The converged heat capacities for each  $P$  are represented in Figure 12. For  $P = 2, 4, 8$ , we take the curves obtained with the highest  $K$ . For  $P = 16$ , we use a large value of  $K$  such that the curve coincides with the curve for  $P = 8$  at high temperatures. We observe a shift of the transition peak, corresponding to sublimation<sup>17</sup>, towards lower temperatures when using  $P > 1$  replicas: it goes from being at around 10K for  $P = 1$  to around 7K for  $P > 1$ . An analogous shift towards lower temperatures is due to NQE and has been observed in Ref. 55 for the 13-atom Neon cluster. Furthermore, as for the harmonic potential, we see that when increasing  $P$ , the curves tend to coincide at a lower temperature for two consecutive values of the number of replicas. As for the harmonic oscillator, there is an increase of the heat capacity at very low  $T$  (we suspect that this increase is caused by using a finite value of  $P$ ). Indeed, this behavior appears at lower temperature for increasing  $P$ , as was the case for the harmonic potential.

To summarize, a system for which NQE are more important — high de Boer parameter, low temperature — requires a larger number of replicas to recover the heat capacity curves. We have also seen that the increase of the number of replicas  $P$  leads to an increase of the number of live points  $K$  resulting in an increase of the computational cost of the exploration. As this procedure must be repeated at each temperature within the direct method, the total amount of computer time can be considerable. In the next section, we will show how it could be reduced by sampling the extended partition function in a single exploration for all temperatures.

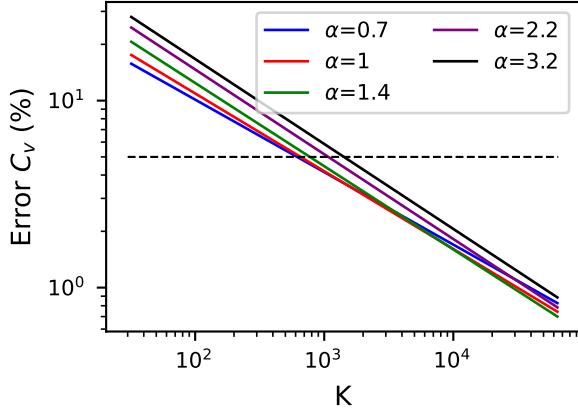


FIG. 13. MAPE error for the 3-atoms Neon cluster with  $P = 2$  for different values of  $\alpha$ . The black dashed line represents an error of 5%. The full lines are the linear log-log fits for each value of  $P$ . We only show the fits obtained from the points for readability. The slopes are  $-0.388$ ,  $-0.416$ ,  $-0.445$ ,  $-0.453$  and  $-0.454$  for  $\alpha = 0.7$ ,  $\alpha = 1$ ,  $\alpha = 1.4$ ,  $\alpha = 2.2$  and  $\alpha = 3.2$ , respectively.

## 2. Extended method

The curves obtained for the highest number of live points ( $K = 8000$  for  $P = 2$ ,  $K = 16000$  for  $P = 4$  and  $K = 32000$  for  $P = 8$ ) with the direct method are considered as a reference for the heat capacity as obtained with the extended method. Here, we do not consider the two lowest temperatures studied using the direct method as they exhibit this increase of the heat capacity (likely due to using a finite value of  $P$ ) and are not converged (in terms of  $P$ ). Moreover, we only adopt the Gaussian window and not the rectangular one as we have seen that the latter leads to larger fluctuations than the former. Furthermore, we use the change of variables presented in Section III B. In that case, choosing the bounds of the parameters is more complex. For Lennard-Jones clusters, the size of the box affects the density and hence the heat capacity curves: the choice of the bounds is discussed in detail in Appendix E.

For the harmonic potential, the optimal value of  $\alpha$  depends on the example studied and number of live points  $K$ . Therefore, first we look at the evolution of the MAPE with  $K$  for different values of  $\alpha$  for  $P = 2$  (Figure 13), starting with the same number of live points considered for the direct method, which we then double in order to check the convergence. We find that, for this range of  $K$  values,  $\alpha = 0.7, 1, 1.4$  give similar errors, which are smaller than for higher values of  $\alpha$ . Similar curves for  $P = 4, 8$  are provided in the supplementary material (Figure 3) Therefore, in the following, we use  $\alpha = 1$  for  $P = 2, 4$  and  $\alpha = 1.4$  for  $P \geq 8$ .

Secondly, we look at the MAPE for  $P = 2, 4, 8$  to find the minimum number of live points required for convergence. We represent the error in Figure 14 for  $P = 2, 4, 8$  to compare the results obtained when using the direct and extended methods. First, we can see that, with the extended method, we need a larger  $K$  to go below the 5% error threshold, which is expected as we sample all temperatures at once. We therefore need more points to sample the energy surface at all temper-

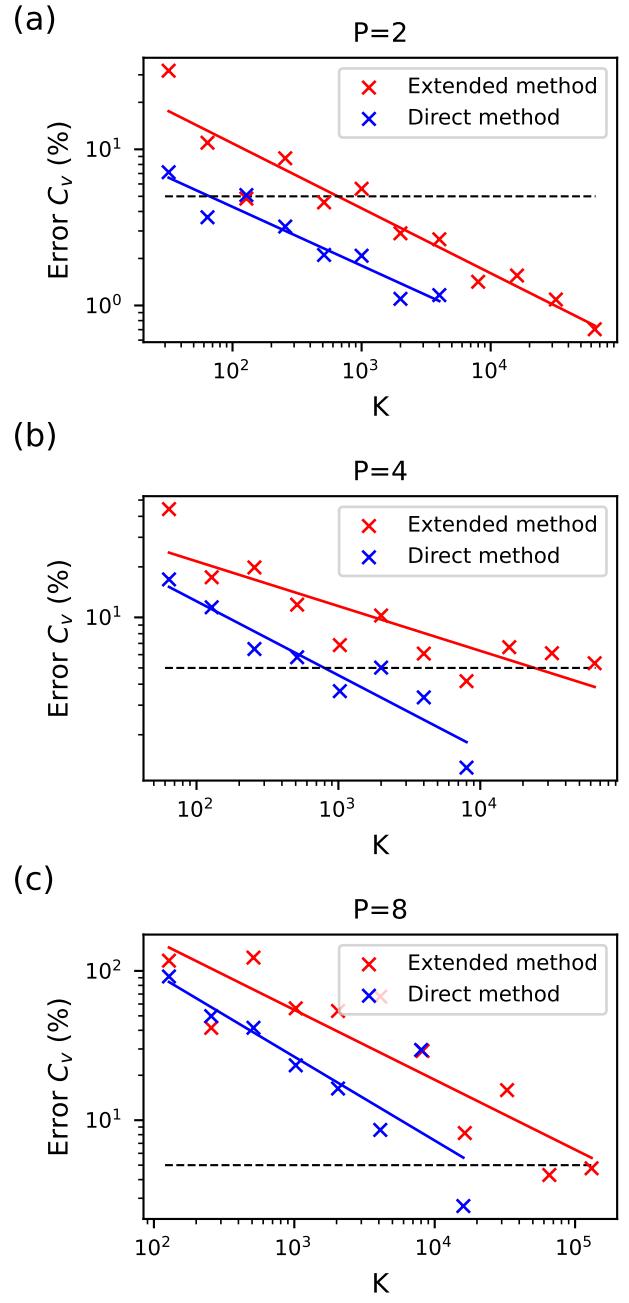


FIG. 14. MAPE error for the 3-atoms Neon cluster with  $P = 2, 4, 8$  with the direct and extended methods. The black dashed line represents an error of 5%. The full lines are the linear log-log fits. We have  $\alpha = 1$  for  $P = 2, 4$  and  $\alpha = 1.4$  for  $P = 8$ . For the extended method, the slopes of the fits are  $-0.416$ ,  $-0.267$  and  $-0.468$  for  $P = 2$ ,  $P = 4$  and  $P = 8$ , respectively. For the direct method, the slopes are indicated in Figure 11.

atures. For the converged value of  $K$ , we again take the first point below the 5% line, that is,  $K = 512$  for  $P = 2$ ,  $K = 8000$  for  $P = 4$  and  $K = 65536$  for  $P = 8$ . We can see that we need 4 to 8 times more live points than with the direct method (for which the values were  $K = 64$ ,  $K = 1000$  and  $K = 16000$ , respectively). In Figure 15, we see that the extended method

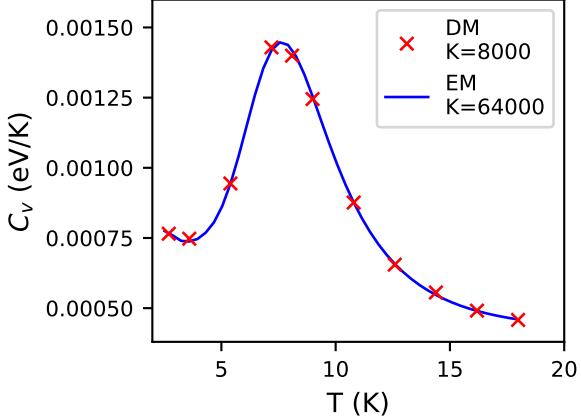


FIG. 15. Comparison of the heat capacity for the 3-atoms Neon cluster between the direct method (DM) and the extended method (EM) for  $P = 2$ . We have  $\alpha = 1$ .

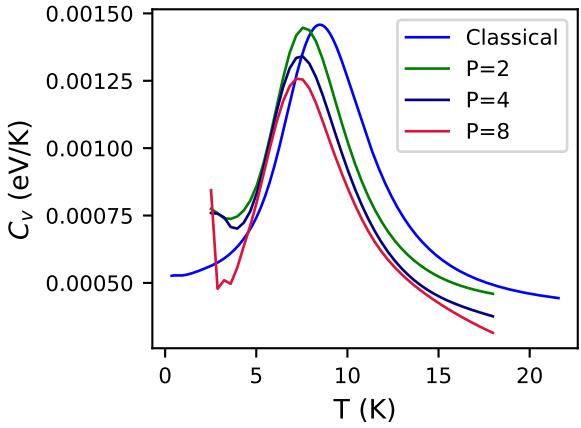


FIG. 16. Heat capacity for the 3-atoms Neon cluster using the extended method. The explorations with  $K = 64000$ ,  $K = 64000$  and  $K = 131072$  are represented for  $P = 2$ ,  $P = 4$  and  $P = 8$ , respectively.

is able to recover the heat capacity curve, as found with the direct method. More details are given in the supplementary material (Figure 4). The two methods therefore converge towards the same result. In Figure 16, the converged heat capacities for each  $P$  are represented as obtained with the extended method using the highest number of live points, similarly to Figure 12 for the direct method. Furthermore, we see that, as for the direct method, the MAPE decreases with the value of  $K$  with a slope of  $\sim -0.5$  in log-log scale, as expected.

Finally, we look at the number of energy evaluations made in the direct and extended methods for  $P = 2, 4, 8$  for the converged explorations. The results are shown in Table I. We can see that to obtain a converged curve, the extended method requires fewer energy evaluations, which is expected as only one exploration is needed. Furthermore, with the direct method, we only access the sampled temperature while the extended

TABLE I. Total number of energy evaluations made by `nested_fit` for the 3-atom Neon cluster for both methods in the converged cases (according to MAPE with a threshold of 5%). In the direct case, 13 explorations were performed for  $P = 2, 4$  and 9 for  $P = 8$ . Twenty cores were used for  $P = 2, 4$  and 64 for  $P = 8$  (see main text).

Number of replicas	Direct method	Extended method
$P = 2$	$2.6 \times 10^7$	$1.4 \times 10^7$
$P = 4$	$1.7 \times 10^9$	$7.7 \times 10^8$
$P = 8$	$5.8 \times 10^{10}$	$3.8 \times 10^{10}$

method provides the thermodynamic properties for all temperature within the studied range. Hence, if we want to compute the thermodynamic properties at a new temperature, it is simply a post-processing step for the extended method while a new exploration must be performed for the direct method.

On this example, we have seen that the extended method requires more live points than the direct method. We have also seen that the optimal value of  $\alpha$  depends on the example studied.

#### D. 13-atom Neon cluster $\text{Ne}_{13}$

We now turn to the 13-atom Neon cluster  $\text{Ne}_{13}$ . This example was previously studied in other works such as Refs. 55–57 using different sampling schemes (Gaussian wavepacket Monte Carlo<sup>55</sup>, path-integral Monte-Carlo<sup>56</sup> and the use of effective potentials and the harmonic superposition approximation<sup>57</sup>). We therefore compare our results using nested sampling with those obtained by Refs. 55 and 56: indeed, both use the same implementation of the Lennard-Jones potential.

In this case, the potential  $V$  in  $V_P$  (Eq. (32)) is the potential presented in Ref. 56, also used in Ref. 55:

$$V(\mathbf{x}) = V_{\text{LJ}}(\mathbf{x}) + \sum_{i=1}^N V_c(\mathbf{x}_i), \quad (57)$$

where  $V_{\text{LJ}}$  is the full Lennard-Jones potential from Eq. (55),  $\mathbf{x}$  the position vector of all  $N$  atoms,  $\mathbf{x}_i$  the position vector of atom  $i$  and  $V_c$  a confining potential which has the following form

$$V_c(\mathbf{x}_i) = \epsilon \left( \frac{|\mathbf{x}_i - \mathbf{x}_c|}{2r_0} \right)^{20} \quad (58)$$

with  $r_0$  and  $\epsilon$  the Lennard-Jones parameters and  $\mathbf{x}_c = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i$  the center of mass of the cluster. The confining potential is used to prevent the atoms from leaving the cluster<sup>56</sup> (in Ref. 55, the confining potential is replaced by the constraint that the initial positions have to be within a distance of  $2r_0$  from the center of mass). We sample the extended partition function with  $\alpha = 1$ . In that case, it is easier to choose the bounds of the parameters than for Lennard-Jones clusters without confining potential: we simply need the sampling space to be large enough so that the confining potential prevents the atoms from leaving the space. Hence, we take  $\bar{\mathbf{x}}$  and

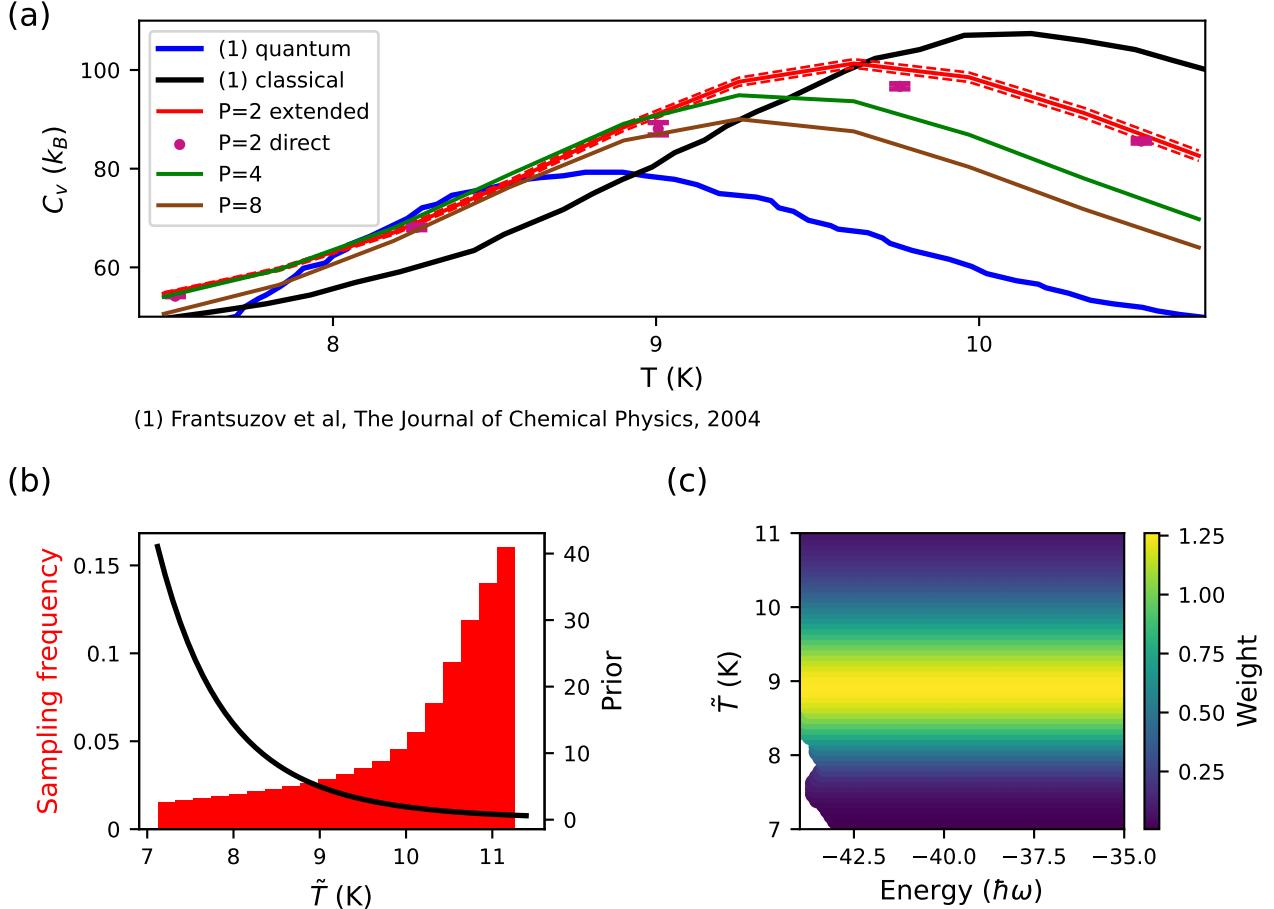


FIG. 17. (a) Heat capacity for the 13-atoms Neon cluster. Comparison with the curve FC-VGW-MC in Ref. 55 using  $\tilde{\beta}^8$  as an extra parameter. The curves for  $P = 2$  and  $P = 4$  are converged in  $K$ . The dashed red lines correspond to the first standard deviation for the  $P = 2$  extended run. For the direct runs, the error bars are smaller than the symbols. The statistical uncertainties were computed from four runs. (b) Sampling frequency of the temperature for the 13-atom Neon cluster as yielded by the extended method using  $\tilde{\beta}^8$  as an extra parameter for  $P = 4$ . We have  $\tilde{T} = 1/(k_B \tilde{\beta})$ . The prior  $\propto \tilde{T}^{-9}$  is indicated by the black curve. (c) Plot of the auxiliary temperature  $\tilde{T}$  against the energy of the sampled configurations for the  $P = 2$  case using  $\tilde{\beta}^8$  as an extra parameter. The points are colored according to the weight given by the Gaussian approximation with  $\alpha = 5$  and  $T = 9$  K.

$\tilde{y}_i$  for  $i = 1, \dots, P - 1$  in a box of side  $6r_0$  centered in 0: we can therefore have  $|\mathbf{x}_i - \mathbf{x}_c| = 6r_0$  which gives  $V_c(\mathbf{x}_i) \approx 3 \times 10^9 \epsilon$ . The classical curve, both from `nested_fit` and Ref. 56 are given in the supplementary material (Figure 5)

We compare our results with the FC-VGW-MC (fully coupled variational Gaussian wave-packet Monte Carlo) method in Ref. 55 that was close to the results found in Ref. 56 — in both cases, the error on the curve is around  $1-2k_B$ . In this case, we put a uniform prior on  $\tilde{\beta}^8$ . As discussed in Section III B 1, the choice of the prior for  $\tilde{\beta}$  has an impact on the sampling distribution of the temperature. There we tested uniform priors on  $\tilde{\beta}^{2s}$ . For  $s < 4$ , low temperatures are poorly sampled (we show the sampling distribution and resulting heat capacity for  $s = 0.5$  in Figure 6 in the supplementary material). For  $s \geq 4$ , low temperatures are sufficiently sampled and no apparent bias appears in the statistical quantities (Figure 17 (b)). In Figure 17 (c), we see that, when computing the thermodynamic properties at 9 K, there is only a small area of signifi-

cant Gaussian weight with no sampled points (corresponding to points with very low energies and low temperatures). This is an improvement to the sampling obtained using a uniform prior on  $\tilde{\beta}$  where we have few points with an auxiliary temperature below 8.5 K at low energies (supplementary material, Figure 6 (c)). Looking to the heat capacity curves (Figure 17 (a)), we see that, with increasing  $P$ , the curves are closer to the quantum one and farther from the classical one. However, the curve of Ref. 55 is not fully recovered by the extended method, even with  $P = 8$ . Higher values of  $P$  are therefore needed. When  $P$  increases, the sampling frequency becomes more unbalanced towards high temperatures. Possibly, a uniform prior over  $\tilde{\beta}^{2s}$  ( $s > 4$ ) will be necessary to sample the space correctly for higher  $P$ . Finally, the statistical uncertainties obtained for the direct and extended methods are of the same order (the extended method uses double the number of live points compared to the direct method).

With this example, we have seen that the prior we give to

the auxiliary temperature greatly impacts the recovery of the expected curve as the distribution of the sampled temperatures depends on it. One therefore wants to choose this prior so that the resulting distribution is flat or samples more the lower temperatures, for which the NQEs are more important. Here, we have tried different priors to obtain a satisfying distribution but this can be quite computationally expensive as multiple runs with different priors are needed. Note that to test different priors, a small number of live points  $K$  can be used as we expect that the sampling frequency of the auxiliary distribution should be similar than with a higher value of  $K$ .

Overall, this trial-and-error methodology has been used throughout this paper to choose the parameters of the extended method: prior, sampling window  $f$  and its parameter  $\alpha$ . This increases the computational cost of the analysis, as multiple choices are studied. This is especially true for the choice of the prior as a new exploration has to be performed for each prior considered. Changing the window or  $\alpha$  only requires to repeat the post-processing step, which is much less computationally costly. In future works, to avoid this extra cost, a more systematic procedure to fine tune these parameters should be developed via an in-depth analysis of their influence on the extended partition function method in other systems.

## VI. CONCLUSION

In this work, we have applied the nested sampling algorithm to evaluate the partition function with temperature-dependent effective potentials. Due to the explicit dependence on the temperature, the straightforward use of nested sampling implies to run distinct simulations for each temperature, which we refer to as the direct method. We have then proposed a method that allows to perform only one exploration for all temperatures, as in the classical case by focusing on the extended partition function. For this method, the partition function is computed over the configurational space as well as over an auxiliary temperature, corresponding to the temperature appearing in the potential. Hence, there is an extra parameter to sample compared to the direct method. However, even if one exploration of the extended partition function is more expensive than the exploration of the original partition function at a single temperature, the former requires less energy evaluations than the latter to compute the different thermodynamic properties.

Specifically to this work, we considered temperature-dependent potentials resulting from the use of the path-integral formalism. The two methods have been compared on two systems: the harmonic potential and Lennard-Jones clusters. The direct method was first tested on the harmonic potential, for which `nested_fit` was able to recover the expected analytical results. Moreover, it was used to study two types of Lennard-Jones clusters:  $\text{Kr}_7$ , which is almost classical, and  $\text{Ne}_3$  and  $\text{Ne}_{13}$ , for which the nuclear quantum effects are important. For the latter, we have studied the convergence of the heat capacity with the number of live points  $K$  and the number of replicas  $P$ . The number of live points  $K$  increases swiftly with the number of replicas  $P$ .

We have seen that the extended method requires more live points than the direct method to reach convergence but has a smaller computational time, mostly due to only one exploration being done instead of several. Furthermore, we have seen that the extended method has additional parameters, such as the size of the regularized Dirac function  $\alpha$  and the choice of the parameter (with uniform prior) to sample, that need to be finely tuned to each case studied. Among these additional parameters, some need only be chosen at the post-processing step ( $\alpha$ ) while others need to be chosen before the nested sampling exploration. In the former case, the same exploration can be used for various values of the parameter. For the latter, a quick exploration can be performed with a reduced number of sampling points to select the optimal values.

In the future, the extended method with nested sampling could be applied to realistic systems as it has the advantage of requiring a single exploration to compute the thermodynamic properties. Furthermore, we envisage strategies for the fine tuning of the parameters of the extended method, which have a non-negligible impact on the recovery of the curves such as the prior distribution of the auxiliary temperature. Indeed, for the moment, we have used a trial-and-error methodology to choose the values of the parameters. Ideally, an automatic tuning of the parameters could be done. To achieve this, a more in-depth analysis of the method (both using exact results on test systems and recursive strategies) is likely needed in the future.

## SUPPLEMENTARY MATERIAL

The supplementary material includes the discrete expressions of the internal energy and heat capacity for a temperature-dependent potential for both the direct and extended methods; the sampling distribution of the temperature  $\tilde{T}$  for the explorations used in Figure 8; the heat capacity curves obtained for different value of  $K$  for the  $\text{Ne}_3$  case for both the direct and extended methods; the evolution of the MAPE with  $K$  for different values of  $\alpha$  for  $P = 4$  and  $P = 8$  for the  $\text{Ne}_3$  case; the classical heat capacity for  $\text{Ne}_{13}$ , and a study of the  $\text{Ne}_{13}$  case using the extended method with a uniform prior on  $\tilde{\beta}$ .

## ACKNOWLEDGMENTS

This work has been realized with the support of the Sorbonne Center for Artificial Intelligence—Sorbonne University—IDEX SUPER 11-IDEX-0004. L.M. thanks Livia Bartók-Pártay, Florent Calvo, Paola Cinella and Tony Lelièvre for their comments. L.M. thanks César Godinho for his work on `nested_fit`. This work was granted access to the HPC resources of TGCC under the allocation A0130906719 made by GENCI.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## AUTHOR DECLARATIONS

### Conflict of interest

The authors have no conflicts to disclose.

### Author contributions

**Philippe Depondt:** Conceptualization (supporting); methodology (supporting); resources (equal); writing – review and editing (equal). **Fabio Finocchi:** Conceptualization (supporting); methodology (supporting); supervision (supporting); resources (equal); writing – review and editing (equal). **Simon Huppert:** Conceptualization (supporting); methodology (supporting); resources (equal); writing – review and editing (equal). **Lune Maillard:** Conceptualization (lead); data curation (lead); formal analysis (lead); investigation (lead); methodology (lead); software (supporting); validation (lead); visualization (lead); writing - original draft preparation (lead). **Thomas Plé:** Conceptualization (supporting); methodology (supporting). **Julien Salomon:** Conceptualization (supporting); methodology (supporting); supervision (supporting); writing – review and editing (equal). **Martino Trassinelli:** Conceptualization (supporting); data curation (supporting); methodology (supporting); software (lead); supervision (lead); writing – review and editing (equal).

## Appendix A: Nested sampling algorithm

### 1. Approaching the density of states

As presented in Section II B, at each iteration of the nested sampling algorithm, we replace the point  $x_{\text{old}}$  by the point  $x_{\text{new}}$  verifying  $V(x_{\text{new}}) < V(x_{\text{old}})$ . Therefore, at each iteration  $\ell$ , we want to uniformly sample the region of space given by the energy constraint  $V(x) < V(x_{\text{old}})$ . We denote  $\rho_\ell$  the cumulative DOS of this region. Since the points are uniformly distributed in the available space, their cumulative densities of states are uniformly distributed over  $[0, 1]^{10,58}$ . Therefore,  $\rho_\ell / \rho_{\ell-1}$  follows a Beta( $K, 1$ ) distribution — which corresponds to the distribution of the outermost value of a set of  $K$  samples following a Uniform( $[0, 1]$ ) distribution. We thus estimate  $\rho_\ell / \rho_{\ell-1}$  by its geometric mean<sup>9,10,59</sup>, i.e.,  $\rho_\ell / \rho_{\ell-1} \approx e^{-1/K}$ . Therefore

$$\rho_\ell \approx e^{-\ell/K}. \quad (\text{A1})$$

We therefore have that  $w_\ell$ , which approximates the DOS, is written

$$w_\ell = \frac{1}{2} (\rho_{\ell-1} - \rho_{\ell+1}) = \frac{1}{2} \left( e^{-\frac{\ell-1}{K}} - e^{-\frac{\ell+1}{K}} \right). \quad (\text{A2})$$

### 2. Searching new points

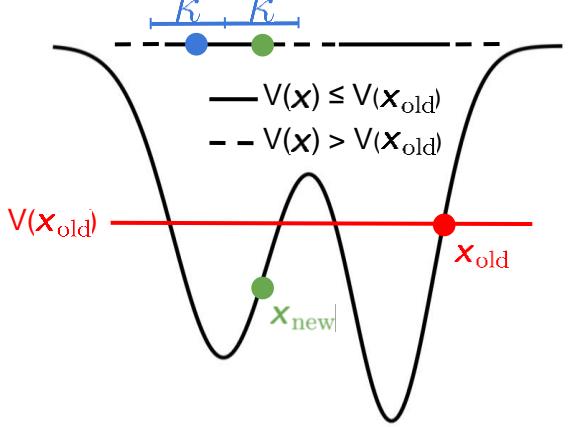


FIG. 18. Representation of slice sampling in 1D. The potential  $V$  is represented in gray and the constraint given by  $V(x_{\text{old}})$  by the red line. The point  $x_{\text{new}}$  is found by taking one of the live points (blue point) and building an interval around it (blue line). A new point verifying the constraint is then sampled (green point).

`Nested_fit`, which is the program used here, uses the slice sampling algorithm<sup>60,61</sup> to find the new point  $x_{\text{new}}$ . This method consists in uniformly choosing new exploration points on a slice of the volume defined by the constraint  $V(x) < V(x_{\text{old}})$ . In the one-dimensional case, represented in Figure 18, one of the live points (blue point in Figure 18) is randomly chosen and an interval, called slice, is built by randomly placing an interval of size  $\kappa$  around it and then extending this interval on both sides — by intervals of size  $\kappa$  — until its limits have an energy higher than  $x_{\text{old}}$  or are out of the sampling space (blue line in Figure 18). A point is then sampled from within the slice and accepted if it satisfies the constraint (green point in Figure 18) and rejected otherwise. In that case, another point is sampled until an acceptable point is found. In a multidimensional setting, a change of coordinates is first performed to efficiently explore all the parameter space, even in presence of strong correlation. This transformation is done via the Cholesky transformation of the covariance matrix of the live points of the considered step to transform the points coordinates into new coordinates with dimensions  $\sim \mathcal{O}(1)$  in all directions<sup>61</sup>. The one-dimensional algorithm is then applied recursively to the vectors of  $n_{\text{bases}}$  randomly generated orthonormal bases<sup>39</sup>. Hence,  $n_{\text{bases}} \times 3N$  steps are performed. The steps can be performed in the transformed space (*Slice Sampling Transformed*) or in the real space (*Slice Sampling Real*)<sup>43</sup>. In this work, we take  $\kappa = 1$  as it satisfies our requirements.

## Appendix B: Discrete expressions for the path-integral internal energy and heat capacity

### 1. Direct method

We denote  $E_i(\beta_j) = V_P(\mathbf{x}^{(i)}, \beta_j)$ ,  $V_i = \bar{V}(\mathbf{x}^{(i)})$  and  $Q_i(\beta_j) = Q(\mathbf{x}^{(i)}, \beta_j)$ , i.e.  $E_i(\beta_j) = V_i + Q_i(\beta_j)$ ,  $\mathbf{x}^{(i)}$  the position vector at iteration  $i$  and  $\beta_j = 1/(k_B T_j)$  the inverse temperature we are considering. From Eqs. (17), (18), (20), (34) and (35), we obtain the following expressions for the internal energy and the heat capacity:

$$U_c^P(\beta_j) = \frac{\sum_i w_i (V_i - Q_i(\beta_j)) e^{-\beta_j E_i(\beta_j)}}{Z_c^P(\beta_j)}, \quad (\text{B1})$$

and

$$\begin{aligned} C_{v,c}^P(\beta_j) &= \frac{\sum_i w_i (V_i - Q_i(\beta_j))^2 e^{-\beta_j E_i(\beta_j)}}{k_B T_j^2 Z_c^P(\beta_j)} \\ &\quad - \frac{\left( \sum_i w_i (V_i - Q_i(\beta_j)) e^{-\beta_j E_i(\beta_j)} \right)^2}{k_B T_j^2 Z_c^P(\beta_j)^2} \\ &\quad - 2 \frac{\sum_i w_i Q_i(\beta_j) e^{-\beta_j E_i(\beta_j)}}{T_j Z_c^P(\beta_j)} \end{aligned} \quad (\text{B2})$$

for an exploration performed at given inverse temperature  $\beta_j$ .

### 2. Extended partition function method

Within the extended partition function method, from Eqs. (27), (28), (26), (34) and (35), the contribution of  $V_P$  to the internal energy and the heat capacity, respectively, read:

$$U_c^P(\beta) \approx \frac{\sum_i w_i (V_i - Q_i) f(\beta - \tilde{\beta}_i; \alpha) e^{-\beta E_i}}{Z_c^P(\beta)} \quad (\text{B3})$$

and

$$\begin{aligned} C_{v,c}^P(\beta) &\approx - \frac{\left( \sum_i w_i (V_i - Q_i) f(\beta - \tilde{\beta}_i; \alpha) e^{-\beta E_i} \right)^2}{k_B T^2 Z_c^P(\beta)^2} \\ &\quad + \frac{\sum_i w_i (V_i - Q_i)^2 f(\beta - \tilde{\beta}_i; \alpha) e^{-\beta E_i}}{k_B T^2 Z_c^P(\beta)} \\ &\quad - 2 \frac{\sum_i w_i Q_i f(\beta - \tilde{\beta}_i; \alpha) e^{-\beta E_i}}{T Z_c^P(\beta)}, \end{aligned} \quad (\text{B4})$$

with  $w_i$  as in Eq. (11) and  $\tilde{\beta}_i$  and  $E_i = V_i + Q_i$  the inverse temperature and energy sampled at iteration  $i$ . The terms  $E_i = E(\mathbf{x}^{(i)}, \tilde{\beta}_i)$  and  $Q_i = Q(\mathbf{x}^{(i)}, \tilde{\beta}_i)$  are computed at inverse temperature  $\tilde{\beta}_i$  for the position vector  $\mathbf{x}^{(i)}$ .

## Appendix C: Derivation of the exact expressions for the path-integral harmonic potential

In this appendix, we derive the exact expressions for the heat capacity of the path-integral harmonic potential with  $P =$

2 and  $P = 4$ . Since all dimensions are separable in Eq. (43), we compute the expressions for one particle in one dimension and multiply the results by 3. We do not use reduced units in this section.

### 1. Exact case with $P = 2$ replicas

Making the change of variables  $y_1 = \mathbf{x}_1 - \mathbf{x}_2$  and  $y_2 = \mathbf{x}_1 + \mathbf{x}_2$ , we have that

$$V_2(\mathbf{x}_1, \mathbf{x}_2, \beta) = \frac{2m}{\hbar^2 \beta^2} (\mathbf{x}_1 - \mathbf{x}_2)^2 + \frac{m\omega^2}{4} (\mathbf{x}_1^2 + \mathbf{x}_2^2) \quad (\text{C1})$$

$$= \left( \frac{2m}{\hbar^2 \beta^2} + \frac{m\omega^2}{8} \right) y_1^2 + \frac{m\omega^2}{8} y_2^2 \quad (\text{C2})$$

$$= V_2(y_1, y_2, \beta). \quad (\text{C3})$$

Therefore, the contribution of  $V_2$  to the partition function (Eq. (37)) is:

$$\begin{aligned} Z_c^{P=2}(\beta) &= \iint_{\mathbb{R}^2} d\mathbf{x}_1 d\mathbf{x}_2 e^{-\beta V_2(\mathbf{x}_1, \mathbf{x}_2, \beta)} \\ &= \frac{1}{2} \iint_{\mathbb{R}^2} dy_1 dy_2 \exp \left( -\beta \left( \left( \frac{2m}{\hbar^2 \beta^2} + \frac{m\omega^2}{8} \right) y_1^2 + \frac{m\omega^2}{8} y_2^2 \right) \right) \end{aligned} \quad (\text{C4})$$

$$= \frac{1}{2} \sqrt{\frac{\pi}{\beta \left( \frac{2m}{\hbar^2 \beta^2} + \frac{m\omega^2}{8} \right)}} \sqrt{\frac{8\pi}{\beta m\omega^2}}. \quad (\text{C5})$$

Consequently, we have

$$C_{v,c}^{P=2}(\beta) = k_B - \frac{4 \frac{(k_B T)^4}{\hbar^4} + \frac{3}{4} \frac{(k_B T)^2 \omega^2}{\hbar^2}}{\left( \frac{(k_B T)^2}{\hbar^2} + \frac{\omega^2}{8} \right)^2} k_B \quad (\text{C7})$$

and

$$C_v^{P=2}(\beta) = 2k_B - \frac{4 \frac{(k_B T)^4}{(\hbar\omega)^4} + \frac{3}{4} \frac{(k_B T)^2}{(\hbar\omega)^2}}{\left( 2 \frac{(k_B T)^2}{(\hbar\omega)^2} + \frac{1}{8} \right)^2} k_B. \quad (\text{C8})$$

### 2. Exact case with $P = 4$ replicas

With the change of variables  $y_1 = \mathbf{x}_1 - \mathbf{x}_2 - \mathbf{x}_3 + \mathbf{x}_4$ ,  $y_2 = \mathbf{x}_1 + \mathbf{x}_2 - \mathbf{x}_3 - \mathbf{x}_4$ ,  $y_3 = \mathbf{x}_1 - \mathbf{x}_2 + \mathbf{x}_3 - \mathbf{x}_4$  and  $y_4 = \mathbf{x}_1 +$

$x_2 + x_3 + x_4$ , we have that

$$\begin{aligned} V_4(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \beta) &= \frac{2m}{\hbar^2 \beta^2} ((\mathbf{x}_1 - \mathbf{x}_2)^2 + (\mathbf{x}_2 - \mathbf{x}_3)^2 + (\mathbf{x}_3 - \mathbf{x}_4)^2 + (\mathbf{x}_4 - \mathbf{x}_1)^2) \\ &\quad + \frac{m\omega^2}{8} (\mathbf{x}_1^2 + \mathbf{x}_2^2 + \mathbf{x}_3^2 + \mathbf{x}_4^2) \end{aligned} \quad (C9)$$

$$\begin{aligned} &= \left( \frac{m}{\hbar^2 \beta^2} + \frac{m\omega^2}{32} \right) y_1^2 + \left( \frac{m}{\hbar^2 \beta^2} + \frac{m\omega^2}{32} \right) y_2^2 \\ &\quad + \left( \frac{2m}{\hbar^2 \beta^2} + \frac{m\omega^2}{32} \right) y_3^2 + \frac{m\omega^2}{32} y_4^2 \end{aligned} \quad (C10)$$

$$= V_4(y_1, y_2, y_3, y_4, \beta). \quad (C11)$$

Hence

$$Z_c^{P=4}(\beta) = C' \sqrt{\frac{1}{\beta^4 \left( \frac{1}{\hbar^2 \beta^2} + \frac{\omega^2}{32} \right) \left( \frac{1}{\hbar^2 \beta^2} + \frac{\omega^2}{32} \right) \left( \frac{2}{\hbar^2 \beta^2} + \frac{\omega^2}{32} \right)}}, \quad (C12)$$

where  $C'$  is a constant that is independent of  $\beta$ . Furthermore,

$$\begin{aligned} C_{v,c}^{P=4}(\beta) &= 2k_B - 2 \frac{\frac{(k_B T)^4}{\hbar^4} + \frac{3}{32} \frac{(k_B T)^2 \omega^2}{\hbar^2}}{\left( \frac{(k_B T)^2}{\hbar^2} + \frac{\omega^2}{32} \right)^2} k_B \\ &\quad - \frac{4 \frac{(k_B T)^4}{\hbar^4} + \frac{3}{16} \frac{(k_B T)^2}{\hbar^2 \omega^2}}{\left( 2 \frac{(k_B T)^2}{\hbar^2} + \frac{\omega^2}{32} \right)^2} k_B \end{aligned} \quad (C13)$$

and

$$\begin{aligned} C_v^{P=4}(\beta) &= 4k_B - 2 \frac{\frac{(k_B T)^4}{(\hbar \omega)^4} + \frac{3}{32} \frac{(k_B T)^2}{(\hbar \omega)^2}}{\left( \frac{(k_B T)^2}{(\hbar \omega)^2} + \frac{1}{32} \right)^2} k_B \\ &\quad - \frac{4 \frac{(k_B T)^4}{(\hbar \omega)^4} + \frac{3}{16} \frac{(k_B T)^2}{(\hbar \omega)^2}}{\left( 2 \frac{(k_B T)^2}{(\hbar \omega)^2} + \frac{1}{32} \right)^2} k_B. \end{aligned} \quad (C14)$$

#### Appendix D: Computational details

The parameters used for the different applications presented in this work are given in Tables II for the harmonic potential with the direct method, III for the harmonic potential with the extended method, IV for Lennard-Jones clusters with the direct method and V for Lennard-Jones clusters with the extended method.

#### Appendix E: Choice of the box size

When presenting the change of variables in Section III B, we mentioned that it makes the choice of the boundaries given to the replicas ( $\tilde{y}_i$ ,  $1 \leq i \leq P-1$ ) in `nested_fit` more complex. There are multiple ways to choose those boundaries, noting  $a$  the size of the interval:

TABLE II. Computational parameters for the quantum harmonic potential with the direct method.

	$K$	$n_{\text{bases}}$	$1/(k_B \beta_s)$	$(k_B T / (\hbar \omega))$
Evolution with $P$		1000	1	0.01
Comparison with extended method		25000	2	0.01

TABLE III. Computational parameters for the quantum harmonic potential with the extended method.

	$K$	$n_{\text{bases}}$	$1/(k_B \beta_s)$	Smeared-delta function + $\alpha$	Explored parameter
Choosing $\alpha$ and $f$	See text	2	0.01	See text	$1/\tilde{\beta}$
Evolution with $P$	See text	2	0.01	Gaussian $\alpha \rightarrow$ see text	See text
Comparison with direct method	50000	2	0.01	Gaussian $\alpha = 7.1$	$1/\tilde{\beta}$

TABLE IV. Computational parameters for the quantum Lennard-Jones clusters with the direct method.

	$K$	$n_{\text{bases}}$	$1/(k_B \beta_s)$	$(k_B T / \epsilon)$
Ne <sub>3</sub>	See text	5	0.01	
Kr <sub>7</sub>	2000 ( $P=1$ ), 5000 ( $P=2$ )	4	0.01	
Ne <sub>13</sub>	16000	3	0.01	

TABLE V. Computational parameters for the quantum Lennard-Jones clusters with the extended method.

	$K$	$n_{\text{bases}}$	$1/(k_B \beta_s)$	Smeared-delta function + $\alpha$	Explored parameter
Ne <sub>3</sub>	See text	2	0.01	Gaussian $\alpha = 1$ ( $P=2,4$ ) $\alpha = 1.4$ ( $P=8,16$ )	$\tilde{\beta}$
Ne <sub>3</sub> (Table I)	See text	5	0.01	/	$\tilde{\beta}$
Ne <sub>13</sub>	See text	3	0.01	Gaussian $\alpha = 1$	$\tilde{\beta}^8$

1. Choosing  $a$  so that it corresponds to the size of the box in real space, i.e.,  $a\lambda_P = L$ .
2. In a similar way, choosing  $a$  so that  $a\lambda_{\text{th}} = L$  where

$$\lambda_{\text{th}} = \sqrt{\frac{2\pi\hbar^2\beta}{m}} \quad (E1)$$

is the de Broglie thermal wavelength which gives information about the delocalization of the particle<sup>62</sup>.

3. If we look at the transformed potential, we have that the contribution to the partition function of the interaction between replicas is  $\exp\left(-\frac{\beta}{2}(\tilde{y}_i - \tilde{y}_{i+1})^2\right)$  which is a

Gaussian function with standard deviation

$$\sigma = \sqrt{\frac{1}{\beta}}. \quad (\text{E2})$$

Hence, we choose  $L$  to be between  $3\sigma$  to  $5\sigma$ .

In the three cases, the parameters depend on the temperature. In the first two cases, the box is bigger for low temperatures and smaller for high temperatures at constant  $a$ : we choose  $a$  at the lowest temperature, so that the replicas always stay in the box at any temperature. For the last case, on the contrary, the box is bigger for high temperatures than for low temperatures; we thus take the box corresponding to the highest temperatures. Figure 19 presents the curves obtained for the  $P = 8$  case using  $\lambda_{th}$  (Eq. (E1)) and  $\lambda_P$  (Eq. (39)) with  $T = 0.01k_B T/\epsilon$ , and 3 and 5 times the Gaussian standard deviation with  $T = 0.35k_B T/\epsilon$ . We clearly see that, when we take the lowest temperature with  $\lambda_P$  and  $\lambda_{th}$ , the heat capacity is not correctly recovered for the higher range of temperatures, the volume explored in that case probably not correctly representing the energy surface. This effect is less visible for lower  $P$  as the parameter space is smaller. Choosing the Gaussian with  $3\sigma$  or  $5\sigma$  recovers the reference curve quite well. The lowest temperature point is of little interest as it is at a non-converged temperature (in terms of  $P$ ). Hence, we use a box of size  $3\sigma$  or  $5\sigma$ . For  $P = 2$  and  $P = 4$ , we use  $\lambda_{th}$  with  $T = 0.01k_B T/\epsilon$ , as this choice only becomes visibly inadequate for  $P \geq 8$ , for which we use a  $3\sigma$  box.

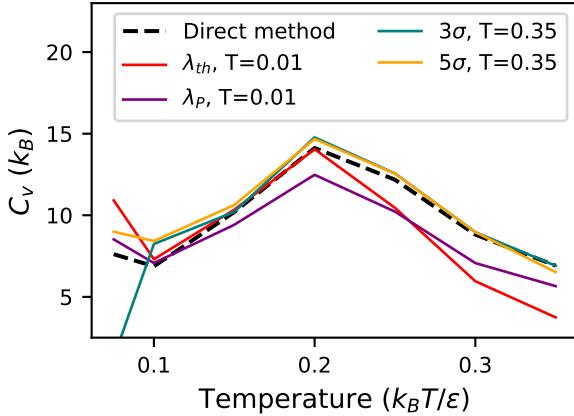


FIG. 19. Heat capacity for the 3-atoms Neon cluster with  $P = 8$  replicas and  $K = 32768$  live points. Comparison of the different choices for the size of the box. The reference curve in dashed blue was obtained with the direct method with the highest  $K$ . We do not consider temperature below  $0.075 \frac{k_B T}{\epsilon}$  as they are not converged with  $P$ . The quantities  $\lambda_{th}$ ,  $\lambda_P$  and  $\sigma$  are defined in Eqs. (E1), (39) and (E2), respectively.

<sup>1</sup>K. Hukushima and K. Nemoto, “Exchange Monte Carlo Method and Application to Spin Glass Simulations,” Journal of the Physical Society of Japan **65**, 1604–1608 (1996).

<sup>2</sup>F. Calvo, J. P. Neirotti, D. L. Freeman, and J. D. Doll, “Phase changes in 38-atom Lennard-Jones clusters. II. A parallel tempering study of equilibrium and dynamic properties in the molecular dynamics and microcanonical ensembles,” The Journal of Chemical Physics **122**, 10350–10357 (2000).

- <sup>3</sup>C. Predescu, P. A. Frantsuzov, and V. A. Mandelshtam, “Thermodynamics and equilibrium structure of Ne38 cluster: Quantum mechanics versus classical,” The Journal of Chemical Physics **122**, 154305 (2005).
- <sup>4</sup>M. Falcioni and M. W. Deem, “A biased Monte Carlo scheme for zeolite structure solution,” The Journal of Chemical Physics **110**, 1754–1766 (1999).
- <sup>5</sup>Y.-H. Lee and B. J. Berne, “Global Optimization: Quantum Thermal Annealing with Path Integral Monte Carlo,” The Journal of Physical Chemistry A **104**, 86–95 (2000).
- <sup>6</sup>G.-F. Hao, W.-F. Xu, S.-G. Yang, and G.-F. Yang, “Multiple Simulated Annealing-Molecular Dynamics (MSA-MD) for Conformational Space Search of Peptide and Miniprotein,” Scientific Reports **5**, 15568 (2015).
- <sup>7</sup>C. Micheletti, A. Laio, and M. Parrinello, “Reconstructing the Density of States by History-Dependent Metadynamics,” Physical Review Letters **92**, 170601 (2004).
- <sup>8</sup>R. Quhe, M. Nava, P. Tiwary, and M. Parrinello, “Path Integral Metadynamics,” Journal of Chemical Theory and Computation **11**, 1383–1388 (2015).
- <sup>9</sup>J. Skilling, “Nested Sampling,” in *AIP Conference Proceedings*, Vol. 735 (AIP, Garching (Germany), 2004) pp. 395–405, iSSN: 0094243X.
- <sup>10</sup>G. Ashton, N. Bernstein, J. Buchner, X. Chen, G. Csányi, A. Fowlie, F. Feroz, M. Griffiths, W. Handley, M. Habeck, E. Higson, M. Hobson, A. Lasenby, D. Parkinson, L. B. Pártay, M. Pitkin, D. Schneider, J. S. Speagle, L. South, J. Veitch, P. Wacker, D. J. Wales, and D. Yallup, “Nested sampling for physical scientists,” Nature Reviews Methods Primers **2**, 39 (2022), arXiv:2205.15570 [astro-ph, physics:cond-mat, physics:hep-ph, stat].
- <sup>11</sup>P. Mukherjee, D. Parkinson, and A. R. Liddle, “A Nested Sampling Algorithm for Cosmological Model Selection,” The Astrophysical Journal **638**, L51 (2006), publisher: IOP Publishing.
- <sup>12</sup>J. Mikelson and M. Khammash, “Likelihood-free nested sampling for parameter inference of biochemical reaction networks,” PLOS Computational Biology **16**, e1008264 (2020).
- <sup>13</sup>S. Das, M. Hobson, F. Feroz, X. Chen, S. Phadke, J. Goudsward, and D. Hohl, “Microseismic event detection in large heterogeneous velocity models using Bayesian multimodal nested sampling,” Data-Centric Engineering **2** (2021), 10.1017/dce.2021.1.
- <sup>14</sup>M. Evans, “Discussion of nested sampling for Bayesian computations by John Skilling,” Bayesian Statistics **8**, 491–524 (2007).
- <sup>15</sup>N. Chopin and C. Robert, “Properties of Nested Sampling,” Biometrika **97**, 741–755 (2010), arXiv:0801.3887 [math, stat].
- <sup>16</sup>I. Murray, D. MacKay, Z. Ghahramani, and J. Skilling, “Nested sampling for Potts models,” in *Advances in Neural Information Processing Systems*, Vol. 18 (MIT Press, 2005).
- <sup>17</sup>L. B. Pártay, A. P. Bartók, and G. Csányi, “Efficient Sampling of Atomic Configurational Spaces,” The Journal of Physical Chemistry B **114**, 10502–10512 (2010).
- <sup>18</sup>M. Griffiths and D. J. Wales, “Nested Basin-Sampling,” Journal of Chemical Theory and Computation **15**, 6865–6881 (2019), publisher: American Chemical Society.
- <sup>19</sup>P. G. Bolhuis and G. Csányi, “Nested Transition Path Sampling,” Physical Review Letters **120**, 250601 (2018).
- <sup>20</sup>G. A. Marchant and L. B. Pártay, “Nested Sampling of Materials’ Potential Energy Surfaces: Case Study of Zirconium,” Physical Sciences Forum **5**, 5 (2022), number: 1 Publisher: Multidisciplinary Digital Publishing Institute.
- <sup>21</sup>B. Szekeres, L. B. Pártay, and E. Mátyus, “Direct Computation of the Quantum Partition Function by Path-Integral Nested Sampling,” Journal of Chemical Theory and Computation **14**, 4353–4359 (2018).
- <sup>22</sup>A. K. Kerman and S. Levit, “Mean-field study of the nuclear partition function: application to level density and compound nucleus fission,” Physical Review C **24**, 1029–1046 (1981).
- <sup>23</sup>N. Goldenfeld, *Lectures on phase transitions and the renormalization group* (CRC Pres, 2018).
- <sup>24</sup>C. Kittel and P. McEuen, *Introduction to solid state physics* (John Wiley & Sons, 2018).
- <sup>25</sup>M. Baggio and A. Zaccone, “Explaining the specific heat of liquids based on instantaneous normal modes,” Physical Review E **104**, 014103 (2021).
- <sup>26</sup>R. P. Feynman, “Space-Time Approach to Non-Relativistic Quantum Mechanics,” Reviews of Modern Physics **20**, 367–387 (1948).
- <sup>27</sup>T. E. Markland and M. Ceriotti, “Nuclear quantum effects enter the main-

stream," *Nature Reviews Chemistry* **2**, 0109 (2018), arXiv:1803.01037 [cond-mat, physics:physics].

<sup>28</sup>M. Ceriotti, W. Fang, P. G. Kusalik, R. H. McKenzie, A. Michaelides, M. A. Morales, and T. E. Markland, "Nuclear Quantum Effects in Water and Aqueous Systems: Experiment, Theory, and Current Challenges," *Chemical Reviews* **116**, 7529–7550 (2016).

<sup>29</sup>J. A. Morrone and R. Car, "Nuclear Quantum Effects in Water," *Physical Review Letters* **101** (2008), 10.1103/physrevlett.101.017801.

<sup>30</sup>K. A. Müller and H. Burkard, "SrTi O 3 : An intrinsic quantum paraelectric below 4 K," *Physical Review B* **19**, 3593–3602 (1979).

<sup>31</sup>S. Schaack, E. Mangaud, E. Fallacara, S. Huppert, P. Depondt, and F. Finocchi, "When Quantum Fluctuations Meet Structural Instabilities: The Isotope- and Pressure-Induced Phase Transition in the Quantum Paraelectric NaOH," *Physical Review Letters* **131**, 126101 (2023), publisher: American Physical Society.

<sup>32</sup>M. Benoit, D. Marx, and M. Parrinello, "Tunnelling and zero-point motion in high-pressure ice," *Nature* **392**, 258–261 (1998).

<sup>33</sup>Y. Bronstein, P. Depondt, L. E. Bove, R. Gaal, A. M. Saitta, and F. Finocchi, "Quantum versus classical protons in pure and salty ice under pressure," *Physical Review B* **93**, 024104 (2016).

<sup>34</sup>S. Schaack, P. Depondt, S. Huppert, and F. Finocchi, "Quantum driven proton diffusion in brucite-like minerals under high pressure," *Scientific Reports* **10**, 8123 (2020).

<sup>35</sup>M. Ceriotti and T. E. Markland, "Efficient methods and practical guidelines for simulating isotope effects," *The Journal of Chemical Physics* **138** (2013), 10.1063/1.4772676, publisher: AIP Publishing.

<sup>36</sup>B. Cheng, E. A. Engel, J. Behler, C. Dellago, and M. Ceriotti, "Ab initio thermodynamics of liquid and solid water," *Proceedings of the National Academy of Sciences* **116**, 1110–1115 (2019).

<sup>37</sup>L. B. Pártay, G. Csányi, and N. Bernstein, "Nested sampling for materials," *The European Physical Journal B* **94**, 159 (2021).

<sup>38</sup>D. Frenkel and B. Smit, *Understanding molecular simulation: from algorithms to applications* (Elsevier, 2023).

<sup>39</sup>L. Maillard, F. Finocchi, and M. Trassinelli, "Assessing Search and Unsupervised Clustering Algorithms in Nested Sampling," *Entropy* **25**, 347 (2023).

<sup>40</sup>M. Trassinelli, "The Nested\_fit Data Analysis Program," in *The 39th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering* (MDPI, 2019) p. 14.

<sup>41</sup>M. Trassinelli, "Bayesian data analysis tools for atomic physics," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **408**, 301–312 (2017).

<sup>42</sup>M. Trassinelli and P. Ciccodicola, "Mean Shift Cluster Recognition Method Implementation in the Nested Sampling Algorithm," *Entropy* **22**, 185 (2020).

<sup>43</sup>L. Maillard, F. Finocchi, C. Godinho, and M. Trassinelli, "Nested sampling for exploring lennard-jones clusters," (2025), arXiv:2501.11370 [physics.comp-ph].

<sup>44</sup>R. P. Kanwal, *Generalized functions: theory and applications* (Springer Science + Business Media, 2004).

<sup>45</sup>R. P. Feynman, A. R. Hibbs, and D. F. Styer, *Quantum mechanics and path integrals* (Courier Corporation, 2010).

<sup>46</sup>M. E. Tuckerman, *Statistical mechanics: theory and molecular simulation* (Oxford University Press, Oxford ; New York, 2010) oCLC: ocn551495372.

<sup>47</sup>R. L. Liboff, *Introductory quantum mechanics* (Pearson Education India, 2003).

<sup>48</sup>P. Goodwin and R. Lawton, "On the asymmetry of the symmetric MAPE," *International Journal of Forecasting* **15**, 405–408 (1999).

<sup>49</sup>J. De Boer and A. Michels, "Contribution to the quantum-mechanical theory of the equation of state and the law of corresponding states. Determination of the law of force of helium," *Physica* **5**, 945–957 (1938).

<sup>50</sup>J. De Boer, "Quantum theory of condensed permanent gases I the law of corresponding states," *Physica* **14**, 139–148 (1948).

<sup>51</sup>M. B. Sevryuk, J. P. Toennies, and D. M. Ceperley, "Why are para-hydrogen clusters superfluid? A quantum theorem of corresponding states study," *The Journal of Chemical Physics* **133**, 064505 (2010).

<sup>52</sup>N. W. Ashcroft and N. D. Mermin, *Solid state physics* (Holt- Rinehart and Winston, New York Chicago San Francisco [etc, 1976]).

<sup>53</sup>This is done as `nested_fit` may encounter problem with the Cholesky transformation of the covariance matrix otherwise.

<sup>54</sup>J. Skilling, P. M. Goggans, and C.-Y. Chan, "Nested Sampling's Convergence," (Oxford (Mississippi), 2009) pp. 277–291.

<sup>55</sup>P. A. Frantsuzov and V. A. Mandelshtam, "Quantum statistical mechanics with Gaussians: Equilibrium properties of van der Waals clusters," *The Journal of Chemical Physics* **121**, 9247–9256 (2004).

<sup>56</sup>C. Predescu, D. Sabo, J. D. Doll, and D. L. Freeman, "Heat capacity estimators for random series path-integral methods by finite-difference schemes," *The Journal of Chemical Physics* **119**, 12119–12128 (2003).

<sup>57</sup>F. Calvo, J. P. K. Doye, and D. J. Wales, "Quantum partition functions from classical distributions: Application to rare-gas clusters," *The Journal of Chemical Physics* **114**, 7312–7329 (2001).

<sup>58</sup>J. S. Speagle, "dynesty: a dynamic nested sampling package for estimating Bayesian posteriors and evidences," *Monthly Notices of the Royal Astronomical Society* **493**, 3132–3158 (2020).

<sup>59</sup>J. Skilling, "Nested sampling for general Bayesian computation," *Bayesian Analysis* **1** (2006), 10.1214/06-BA127.

<sup>60</sup>R. M. Neal, "Slice sampling," *The Annals of Statistics* **31** (2003), 10.1214/aos/1056562461.

<sup>61</sup>W. J. Handley, M. P. Hobson, and A. N. Lasenby, "polychord: next-generation nested sampling," *Monthly Notices of the Royal Astronomical Society* **453**, 4385–4399 (2015).

<sup>62</sup>I. F. Silvera, "Bose–Einstein condensation," *American Journal of Physics* **65**, 570–574 (1997).