Nested Sampling for Exploring Lennard-Jones Clusters

Lune Maillard, Fabio Finocchi, César Godinho, Martino Trassinelli

Nested Sampling for Exploring Lennard - Jones Clusters

Lune Maillard[1], Fabio Finocchi[1], César Godinho[2]
and Martino Trassinelli[1]
[1] Institut des Nanosciences de Paris, Sorbonne Université, CNRS, France
[2] Laboratory of Instrumentation, Biomedical Engineering and Radiation Physics (LIBPhys-UNL), Department of Physics, NOVA School of Science and Technology, NOVA University Lisbon, Caparica, Portugal
E-mail: lune.maillard@insp.upmc.fr (L.M.), martino.trassinelli@insp.jussieu.fr (M.T)

Abstract

Lennard - Jones clusters, while an easy system, have a significant number of non equivalent configurations that increases rapidly with the number of atoms in the cluster. Here, we aim at determining the cluster partition function; we use the nested sampling algorithm, which transforms the multidimensional integral into a one-dimensional one, to perform this task. In particular, we use the `nested_fit` program, which implements slice sampling as search algorithm. We study here the 7-atoms and 36-atoms clusters and then look at the computational cost of `nested_fit` with and without parallelisation. We find that `nested_fit` is able to recover phase transitions and find different stable configurations of the cluster. Furthermore, the implementation of the slice sampling algorithm has a clear impact on the computational cost.

Keywords

Lennard - Jones; Clusters; Nested sampling

1 Introduction

Lennard - Jones clusters are relevant models for rare gas atoms clusters [1] but the number of configurations increases exponentially with the number of degrees of freedom [2]. They have been widely studied using a variety of methods such as basin-hopping [3], Metropolis Monte-Carlo [4], Molecular Dynamics [5] or parallel tempering [6, 7].

In this work, we use the nested sampling algorithm [8] to study those clusters. This algorithm is mainly employed to analyse data in fields such as cosmology and astrophysics [9] but is applied more and more in materials science for the exploration of potential energy surfaces, including Lennard - Jones clusters [10]. Here, we utilise a different implementation of nested sampling than the one used in Ref. [10].

We first present the implementation of the nested sampling algorithm for the specific case of Lennard - Jones cluster in Section 2. In Section 3, we study, in particular, clusters of two sizes. In Section 4, we look at the computational cost of our implementation. Finally, we conclude in Section 5.

2 Nested sampling for Lennard - Jones clusters

For a cluster of N atoms, the truncated and shifted Lennard - Jones potential takes the following form [11, 12]

$$V = \sum_{1 \leq i < j \leq N} \tilde{V}(r_{ij})$$

with

$$\tilde{V}(r) = \begin{cases} V_{LJ}(r) - V_{LJ}(r_c) & r < r_c, \\ 0 & r \geq r_c, \end{cases}$$

and

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

We have that r_{ij} is the distance between atoms i and j and that r_c is the cutoff radius, used to remove the interaction at infinite range. The partition function for this system can be written as

$$\begin{aligned} Z(\beta) &= \frac{1}{h^3} \int \exp\left(-\beta \left(\frac{1}{2} \sum_{i=1}^{3N} \frac{p_i^2}{m} + V(\boldsymbol{x}) \right) \right) d\boldsymbol{x} d\boldsymbol{p} \\ &= \frac{1}{h^3} \int \exp\left(-\beta \left(\frac{1}{2} \sum_{i=1}^{3N} \frac{p_i^2}{m} \right) \right) d\boldsymbol{p} \int \exp(-\beta(V(\boldsymbol{x})) d\boldsymbol{x} \\ &= Z_k * Z_c, \end{aligned}$$

where $\boldsymbol{p} = (p_1, ..., p_{3N})$ is the vector of momenta, $\boldsymbol{x}$ the vector of all atomic positions, $\beta = \frac{1}{k_B T}$ the inverse temperature and h the Planck constant. The separation in the equation above is possible as the potential V only depends on the atom positions and not on the momenta. The term $Z_k = \frac{1}{h^3} \int \exp\left(-\beta \left(\frac{1}{2} \sum_{i=1}^{3N} \frac{p_i^2}{m} \right) \right) d\boldsymbol{p}$ is easy to compute and gives $(\sqrt{2\pi m/(\beta h^2)})^{3N}$. The term $Z_c = \int \exp(-\beta(V(\boldsymbol{x})) d\boldsymbol{x}$ can be rewritten as $Z_c = \int \rho(E) \exp(-\beta E) dE$ where $\rho(E)$ is the density of states [13] and does not in general have an analytical form. We compute Z_c by using a Monte Carlo sampling and, more precisely, the nested sampling algorithm which works in the following way [8, 10]:

1. K points, called live points, are uniformly sampled from the entire space.
2. At each iteration i, the point $\boldsymbol{x}_{old}$ associated to the highest energy is removed and replaced by a point $\boldsymbol{x}_{new}$ with an energy that is strictly lower: $V(\boldsymbol{x}_{new}) < V(\boldsymbol{x}_{old})$. The method to find this new point will be presented in Section 2.1. Denoting $E_i = V(\boldsymbol{x}_{old})$, this point will contribute to the partition function as

$$c_i = w_i e^{-\beta E_i} \text{ with } w_i = \frac{1}{2} \left(\left(\frac{K}{K+1} \right)^{m-1} - \left(\frac{K}{K+1} \right)^{m+1} \right). \quad (1)$$

The term w_i is the approximation of the density of states (DOS) that is evaluated by statistical considerations.

3. This procedure is repeated until the current contribution is small compared to previous contributions: $\log(c_i) - \log(c_{max}) < \delta$ with $c_{max} = \max_{m \leq i}(c_i)$ at an inverse temperature β chosen by the user. Here, we use $\delta = -10$, a value that satisfies our requirements.

As the interaction potential does not depend on the temperature, we can compute the partition function at all temperatures by estimating pairs of (E_i,w_i) values during one single exploration of the potential [10] as

$$Z_c(\beta) \approx \sum_i c_i.$$

From the partition function, other properties of the system can be calculated such as the internal energy

$$U = -\frac{\partial \log(Z)}{\partial \beta} = -\frac{\partial \log(Z_k)}{\partial \beta} - \frac{\partial \log(Z_c)}{\partial \beta} = \frac{3N}{2} k_B T - \frac{\partial \log(Z_c)}{\partial \beta}$$

and the heat capacity

$$C_v = \frac{\partial U}{\partial T} = \frac{3N}{2} k_B - \frac{\partial}{\partial T} \frac{\partial \log(Z_c)}{\partial \beta}.$$

In this paper, we use, in particular, the `nested_fit` program [14–17]. In the next sections, we present some of the specific implementation of the nested sampling algorithm in `nested_fit`.

2.1 Slice sampling and slice sampling transformed

The method used to find the new point $\boldsymbol{x}_{new}$ in `nested_fit` is slice sampling [18, 19]. This method consists in uniformly choosing new exploration points on a slice of the volume defined by the constraint $V(\boldsymbol{x}) < V(\boldsymbol{x}_{old})$. In one dimension, one of the live points is randomly chosen and a slice is built around it until the end points of the slice have an energy higher than $\boldsymbol{x}_{old}$ or are out of the sampling space. A point is then sampled from within the slice and accepted if it verifies the constraint 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 [19]. The one-dimensional algorithm is then applied recursively to the vectors of n_{bases} randomly generated orthonormal bases (see Ref. [17]). Hence, $n_{bases} \times 3N$ steps are performed.

There are two ways of implementing this algorithm:

1. First, the steps can be performed in the transformed space. This will be referred to as *slice sampling transformed* and was used in Ref. [17]. However, the sampled

points need to be transformed back to the real space to compute the energy and check that the points are within the bounds. This brings a significant computational overhead to the computation.

2. Second, the steps can be performed in the real space and only the slices are chosen in the transformed space. This will be referred to as *slice sampling* and was used in Refs. [19, 20]. In that case, only the orthonormal bases need to be transformed to the real space.

2.2 Parallelisation

The nested sampling algorithm can be easily parallelised. One method that is used to make nested sampling parallel is, instead of searching and substituting one point per iteration, to search r new points in parallel and then to replace r points at once [21]: the r points with the highest energy are removed and replaced by r new points with an energy lower than the energies of all removed points. In that case, the compression factor is $\frac{K-r+1}{K}$ instead of $\frac{K}{K+1}$ (see Eq. (1)) [13]. However, this parallelisation increases the variance of the results [13]. A different approach is used in `Polychord` using MPI protocols [19, 20]: one primary process is in charge of the sequential replacing of the live points while all other secondary processes are in charge of finding new live points continuously, with the threshold criterion changed by the primary process. The sequential replacing is thus done by the primary process with the points that are found by the secondary processes. A point found for an iteration with a specific energy constraint is valid for a further iteration if it verifies the new constraint [19]. In this case, the compression factor is still $\frac{K}{K+1}$.

The method we use is very similar to the one used in `Polychord`. First, the point with highest energy $\boldsymbol{x}_{max}$ is identified; we note its associated energy V_{max}. Then, r searches are done in parallel to find a new point with energy lower than V_{max}. We thus obtain a list of r new points each with its associated energy. Finally, the points are sequentially added (without parallelisation) to the set of live points by comparing their energy with V_{max} which evolves with the addition of the new points. There are two possibilities:

- If the energy is lower than V_{max}, the point is added to the live points, $\boldsymbol{x}_{max}$ is removed and V_{max} is updated to the new maximum energy of the set of live points. The new value will be used for new points not yet added or rejected.

- If the energy is higher than V_{max}, the new point is rejected. In this case, V_{max} is not updated.

Hence, for each set of searches done, between one and r points are replaced in our set of live points. Indeed, by definition, all the searches have an energy lower than the initial V_{max} so that at least one point is added.

2.3 Computing the covariance matrix

Another aspect of `nested_fit` that brings a computational overhead is the computation of the covariance matrix and its Cholesky transform at every iteration. To accelerate the program, we have chosen to compute those two matrices only every $0.05K$ iterations. This value was chosen as a trade-off between accuracy and computational cost.

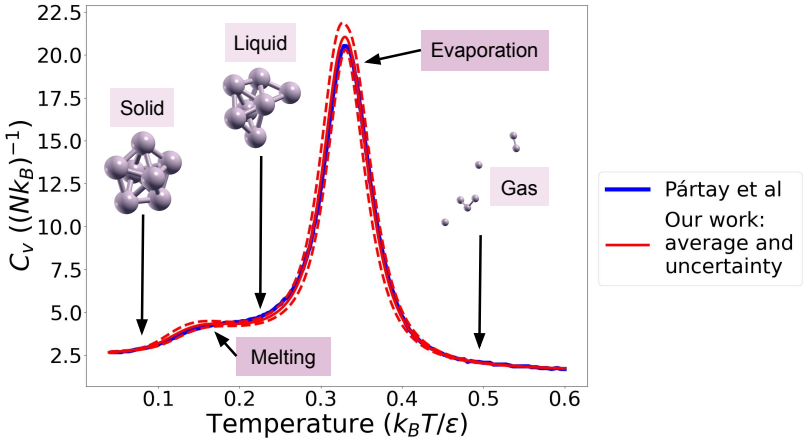

Figure 1: Heat capacity for the Lennard - Jones cluster with 7 atoms. Comparison with the results in [10]. The configurations shown were obtained with `XCrySDen` [24]. We used $K = 1000$ live points and $n_{bases} = 5$. Eight runs were performed.

3 Lennard - Jones clusters

We now study Lennard - Jones clusters of two sizes : 7 atoms and 36 atoms. In the case of the 7-atoms cluster, we compare our results with those found in Ref. [10]. In this work, we use reduced units i.e. the temperature is in units of $\frac{k_BT}{\epsilon}$ with $\epsilon = 1$. In reduced units, the stopping temperature is going to be 0.01 for the 7-atoms cluster and 0.005 for the 36-atoms cluster. For the simulation, the atoms are confined in a cubic box of side L. To fix the size of the box, we will look at the value of the density ρ in unit of r_0^{-3}: $\rho = \frac{N}{(L/r_0)^3} r_0^{-3}$. In all our cases, we take $L = 6$ and use r_0 to tune the density to a given desired value. Within reduced units, the results are valid for all atomic species interacting via the Lennard - Jones potential: the specific (r_0,ϵ) values, corresponding to the atomic species considered, need to be reintroduced instead of the reduced ones. Hence, the specific value of r_0 is not important. The density characterises the size of the space the particles evolve in, relative to the number of atoms N.

Here, we aim to compute the partition function and derivatives of the clusters, aiming at detecting the possible phase transitions. In a infinite system, a phase transition is indicated by a discontinuity in the heat capacity curve, either in the form of a singularity (delta function) or a critical exponent [22]. However, here we work in a finite system composed of N particles in which case a phase transition is not shown by a discontinuity but rather by a peak in the heat capacity curve [23]. This peak can sometimes take the form of a shoulder.

3.1 7 atoms

First, we consider a cluster made of $N = 7$ atoms with a density of $\rho = 9.27 \times 10^{-3} r_0^{-3}$ (corresponding to $r_0 = 0.659$). The size of the cluster and the density are chosen in order to compare our results with those obtained in Ref. [10]. We use $K = 1000$ live points repeating 8 times the procedure to estimate the uncertainties. Ref. [10] uses the `pymatnest` [10, 13, 25, 26] program with rejection Gibbs sampling to update iteratively

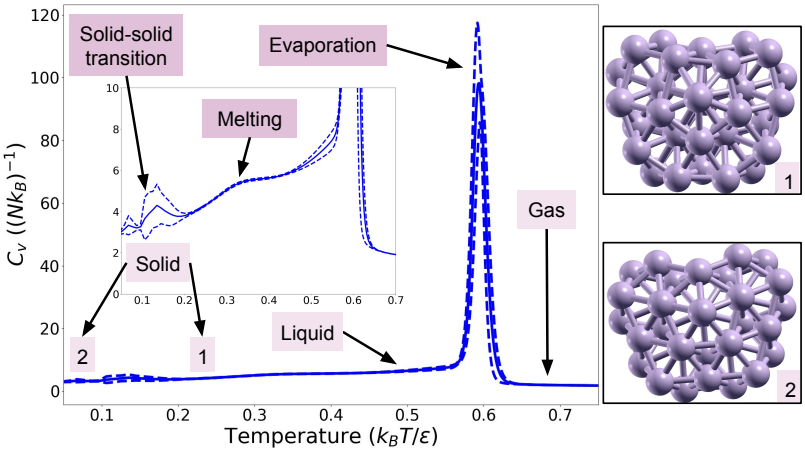

Figure 2: Heat capacity for the Lennard - Jones cluster with 36 atoms. The configurations shown were obtained with `XCrySDen` [24]. We used $K = 70000$ live points and $n_{bases} = 3$. Eight runs were performed. The full line is the mean obtained from the eight runs and the dashed curves correspond to the first standard deviation.

the points — which moves one atom at a time — with $K = 500$ live points. We both use a cutoff radius r_c of $3r_0$. Both heat capacity curves obtained are represented in Figure 1 with examples of solid, liquid and gas configurations. We see that a good agreement with Ref. [10] is obtained. More importantly, `nested_fit` is able to recover both the evaporation and melting phase transitions. We also ran `nested_fit` with $K = 500$ live points but obtained an evaporation peak that was slightly shifted towards higher temperatures compared to the curve from Ref. [10], indicating that there are too few points to obtain convergence. The difference in search method may explain why we need to double the points compared to Ref. [10]. Indeed, `pymatnest` was developed specifically for the exploration of energy surfaces while `nested_fit` has applications in both data analysis and materials science. The search method used by the latter thus have to work in both cases. Furthermore, in Ref. [10], the Lennard - Jones function was called 2×10^7 times while here it was called 5.4×10^7 times. There are two main reasons for this increase: first, we use twice as many live points, which should more or less double the number of calls, and second, in slice sampling, the function is called not only to move the points but also to extend the slice.

3.2 36 atoms

Second, we consider a cluster composed of $N = 36$ atoms with a density of $\rho = 1.79 \times 10^{-2} r_0^{-3}$ (corresponding to $r_0 = 0.475$). This example is quite challenging as a low temperature peak is observed, associated to a Mackay - anti-Mackay phase transition [27, 28]. The heat capacity obtained with $K = 70000$ is shown in Figure 2. As for the previous case, the calculation is repeated 8 times for the estimation of the uncertainties. Again, we can see the presence of a peak corresponding to evaporation and a shoulder corresponding to melting. However, in this case, there is also a peak at lower temperature (around $T = 0.135$) that corresponds to a solid-solid phase transition. From the standard deviation of the heat capacity curve, we can see that the position of this peak is more

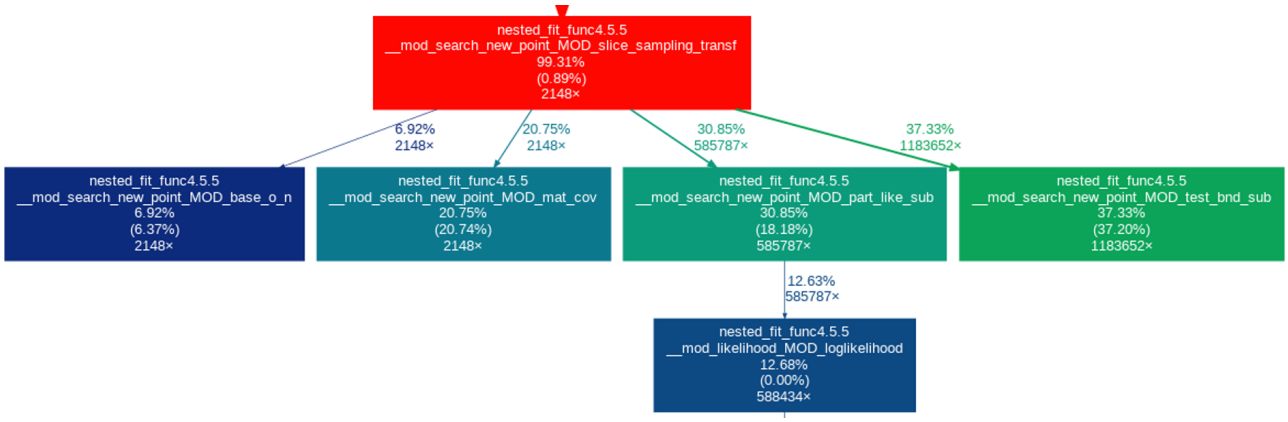

Figure 3: Profiling for slice sampling transformed with the calculation of the covariance matrix at each iteration

unstable from one run to the other than the position of the evaporation peak. Indeed, this low temperature peak is harder to obtain — and was not seen in runs with fewer live points — as we need enough points to fall in each basin so that it is correctly explored. A variation in the proportion of samples in each basin could explain this difference between the runs. We found the positions of the peaks to be similar when more live points - up to $K = 110000$ - were used, indicating that $K = 70000$ live points is enough to attain convergence. The two solid configurations are shown in Figure 2. The low temperature peak was also found in [10] at $T = 0.145$ which is in the interval $[0.1, 0.15]$ given by our eight runs.

4 Program profiling and optimisation

In this section, we look at the impact of the different developments made to `nested_fit` on the computational resources needed. For that, we take the 36-atoms cluster for its high number of parameters (108). First, we look at the result of profiling — performed with `valgrind` and `gprof2dot` —, to see which parts of the algorithm are the most computationally expensive. We used $K = 500$ and $n_{bases} = 1$. We consider three cases: slice sampling transformed with the calculation of the covariance matrix at each iteration, slice sampling transformed with the calculation of the covariance matrix every $0.05K$ iterations and slice sampling with the calculation of the covariance matrix every $0.05K$ iterations. The results of the profiling are presented in Figures 3, 4 and 5 respectively. In the first case, version 4.5.5 of `nested_fit` was used and version 4.6.0 for the two other cases. The purpose of each function is explained in Table 1. Only the relevant boxes are represented. The last three lines in a box corresponds to (1) the percentage of time spent in this function and its children function, (2) the percentage of time spent in this function alone and (3) the number of times the function was called. The arrows linking one box to another represent the first (parent) function calling the second (child) with the two numbers being the percentage of time the child function transfers to its parent and the number of times the parent function calls the child function. A box is only represented

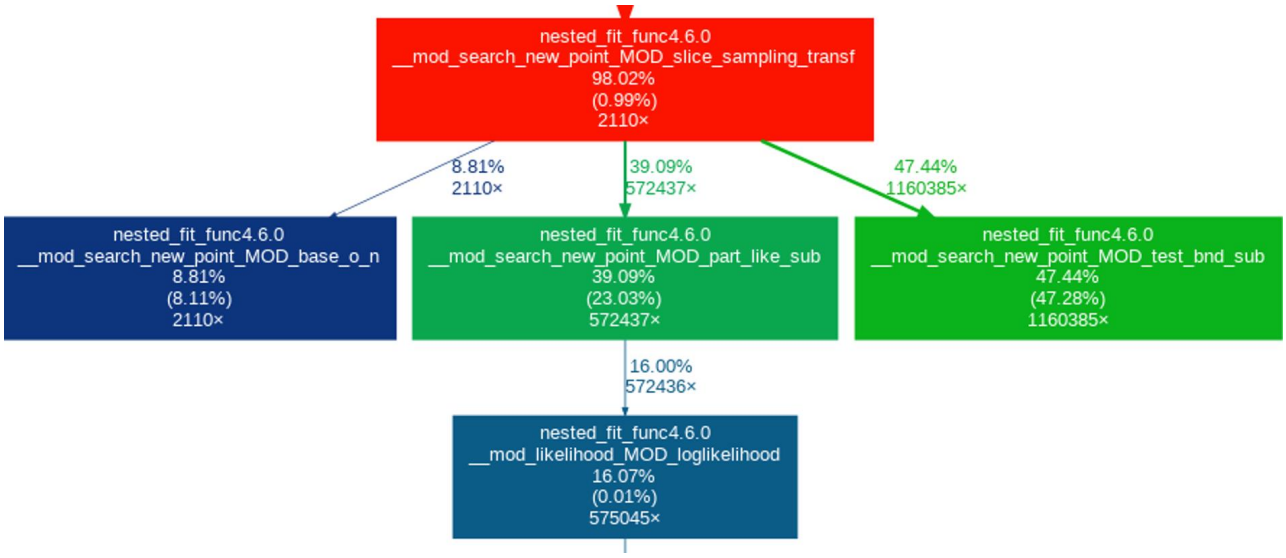

Figure 4: Profiling for slice sampling transformed with the calculation of the covariance matrix every $0.05K$ iterations

if at least 5% of the time is spent in the function.

First, looking at Figures 3 and 4, we can see that, in the case of slice sampling transformed, over 50% of the time is spent going back to the real space to calculate the likelihood and check the boundaries (i.e. in the part_like_sub and test_bnd_sub functions) which is not needed in slice sampling. Thus, more time is spent changing space than calculating the function studied, leading to a computational overhead for slice sampling transformed. Furthermore, we can see that when the covariance function is computed at each iteration (Figure 3), this calculation takes around 20% of the time while when it is only computed every $0.05K$ iterations (here every 25 iterations) (Figures 4 and 5), the box does not appear, which means that less than 5% of the time is spent on this particular task. Finally, we see that, for slice sampling transformed with the covariance matrix computed at each iteration (Figure 3), around 12% of the time is spent calculating the function studied. This number goes up to around 55% (4.4 times more) when using slice sampling and computing the covariance function every $0.05K$ iterations (Figure 5). In both cases, the number of times the energy function was called is roughly the same (around 580000 times). This shows that, with our optimisations, we removed calculations that are not needed and that bring a significant computational overhead. This is further shown by Table 2 that gives the computational time of all three cases for 20000 iterations: with the optimisations, the computational time is reduced by a factor 2.8.

For the impact of the parallelisation on the computation time, we also look at the time required to run `nested_fit` with and without parallelisation using slice sampling transformed with $K = 1000$ and $n_{bases} = 1$. The covariance matrix is computed every iteration. The parallelised case over 64 cores is around 21 times faster than the non-parallelised case (the non-parallelised case takes 85 minutes while the parallelised case takes 4 minutes). A possible explanation for this dividing factor being smaller than the number of cores used is that the time gained by simultaneously searching 64 new points

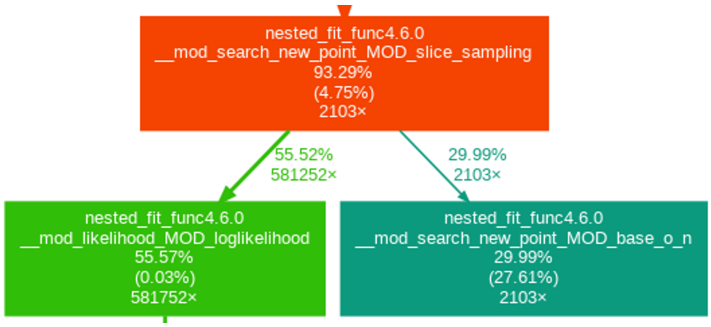

Figure 5: Profiling for slice sampling with the calculation of the covariance matrix every $0.05K$ iterations

__mod__search__new__point__MOD__slice__sampling Performs one iteration of nested sampling with slice sampling
__mod__search__new__point__MOD__slice__sampling__transf Performs one iteration of nested sampling with slice sampling transformed
__mod__likelihood__MOD__loglikelihood Calls the function to calculate (here the Lennard - Jones potential)
__mod__search__new__point__MOD__mat__cov Computes the covariance matrix
__mod__search__new__point__MOD__base__o__n Creates an orthonormal basis
__mod__search__new__point__MOD__part__like__sub Calculates the function for a point in the transformed space
__mod__search__new__point__MOD__test__bnd__sub Checks if a point in the transformed space is inside the sampling space

Table 1: Purpose of the functions shown in Figures 3 to 5.

may be partially counterbalanced by the time taken for the sequential adding of all the points, some of which may be rejected due to the constraint changing with each added point.

5 Conclusion

In this work, we have applied the nested sampling algorithm to study classical Lennard - Jones clusters.

	Version 4.5.5 Slice sampling transformed	Version 4.6.0 Slice sampling transformed	Version 4.6.0 Slice sampling
Time (s)	442.19	367.57	158.13

Table 2: Time taken by all three cases to perform 20000 iterations without parallelisation.

First, we have studied clusters of two sizes. As for the 7-atoms cluster, we have seen that we were able to recover phase transitions by comparing our work with that in Ref. [10]. We saw that we required more live points and more likelihood calls that in Ref. [10], likely due to the different choices of method to find a new point, which is an inconvenient for a more computationally expensive energy function. We then studied the 36-atoms cluster which presents a solid-solid phase transition at low temperature that `nested_fit` was able to recover provided enough live points are used.

We have then looked at the impact of different implementation on the computational resources: we have seen that using slice sampling transformed brought a consequent computational overhead through the need to transform back to the real space to compute the energy. Slice sampling, which works directly in the real space, does not need to perform this task and therefore spends more than triple the time calculating the energy compared to slice sampling transformed, even though the function was called around the same number of times. We also saw that, using parallelisation on 64 cores, we gained a factor 21 in the computation time. All of those developments allowed us to study bigger and more complex clusters that require a higher number of live points. For example, one run for the 36-atoms Lennard - Jones clusters with $K = 70000$ took around 42 hours to run with the parallelisation.

In the future, we would like to study quantum Lennard - Jones clusters which are by about an order of magnitude more computationally expensive than the classical ones.

Acknowledgments

Work realised with the support of the Sorbonne Center for Artificial Intelligence 579 - Sorbonne University - IDEX SUPER 11-IDEX-0004. The authors thank Philipe Depondt, Simon Huppert and Julien Salomon for helpful discussions.

References

[1] N. W. Ashcroft and N. D. Mermin. *Solid state physics*. Brooks/Cole, 1976.

[2] J. P. K. Doye, M. A. Miller, and D. J. Wales. Evolution of the potential energy surface with size for Lennard-Jones clusters. *The Journal of Chemical Physics*, 111(18):8417–8428, 1999.

[3] D. J. Wales and J. P. K. Doye. Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms. *The Journal of Physical Chemistry A*, 101(28):5111–5116, 1997.

[4] N. Quirke and P. Sheng. The melting behavior of small clusters of atoms. *Chemical Physics Letters*, 110(1):63–66, 1984.

[5] T. L. Beck and R. S. Berry. The interplay of structure and dynamics in the melting of small clusters. *The Journal of Chemical Physics*, 88(6):3910–3922, 1988.

[6] J. P. Neirotti, F. Calvo, D. L. Freeman, and J. D. Doll. Phase changes in 38-atom Lennard-Jones clusters. I. A parallel tempering study in the canonical ensemble. *The Journal of Chemical Physics*, 112(23):10340–10349, 2000.

[7] 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*, 112(23):10350–10357, 2000.

[8] J. Skilling. Nested Sampling. In *AIP Conference Proceedings*, volume 735, pages 395–405, Garching (Germany), 2004. AIP. ISSN: 0094243X.

[9] 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(1):39, 2022. arXiv:2205.15570.

[10] L. B. Pártay, A. P. Bartók, and G. Csányi. Efficient Sampling of Atomic Configurational Spaces. *The Journal of Physical Chemistry B*, 114(32):10502–10512, 2010.

[11] W. Shi and J. K. Johnson. Histogram reweighting and finite-size scaling study of the Lennard–Jones fluids. *Fluid Phase Equilibria*, 187-188:171–191, 2001.

[12] L. B. Pártay, C. Ortner, A. P. Bartók, C. J. Pickard, and G. Csányi. Polytypism in the ground state structure of the Lennard-Jonesium. *Physical Chemistry Chemical Physics*, 19(29):19369–19376, 2017.

[13] L. B. Pártay, G. Csányi, and N. Bernstein. Nested sampling for materials. *The European Physical Journal B*, 94(8):159, 2021.

[14] M. Trassinelli. The Nested_fit Data Analysis Program. In *The 39th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, page 14. MDPI, 2019.

[15] 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.

[16] M. Trassinelli and P. Ciccodicola. Mean Shift Cluster Recognition Method Implementation in the Nested Sampling Algorithm. *Entropy*, 22(2):185, 2020.

[17] L. Maillard, F. Finocchi, and M. Trassinelli. Assessing Search and Unsupervised Clustering Algorithms in Nested Sampling. *Entropy*, 25(2):347, 2023.

[18] R. M. Neal. Slice sampling. *The Annals of Statistics*, 31(3), 2003.

[19] W. J. Handley, M. P. Hobson, and A. N. Lasenby. PolyChord: nested sampling for cosmology. *Monthly Notices of the Royal Astronomical Society: Letters*, 450(1):L61–L65, 2015. arXiv:1502.01856.

[20] W. J. Handley, M. P. Hobson, and A. N. Lasenby. polychord: next-generation nested sampling. *Monthly Notices of the Royal Astronomical Society*, 453(4):4385–4399, 2015.

[21] R. W. Henderson and P. M. Goggans. Parallelized nested sampling. *AIP Conference Proceedings*, pages 100–105, 2014.

[22] A. Hüller. First order phase transitions in the canonical and the microcanonical ensemble. *Zeitschrift für Physik B Condensed Matter*, 93(3):401–405, 1994.

[23] I. Medved', A. Trník, and L. Vozár. Modeling of heat capacity peaks and enthalpy jumps of phase-change materials used for thermal energy storage. *International Journal of Heat and Mass Transfer*, 107:123–132, 2017.

[24] A. Kokalj. XCrySDen—a new program for displaying crystalline structures and electron densities. *Journal of Molecular Graphics and Modelling*, 17(3):176–179, 1999.

[25] R. J. N. Baldock, L. B. Pártay, A. P. Bartók, M. C. Payne, and G. Csányi. Determining pressure-temperature phase diagrams of materials. *Physical Review B*, 93(17):174108, 2016. arXiv:1503.03404 [cond-mat, physics:physics].

[26] R. J. N. Baldock, N. Bernstein, K. M. Salerno, L. B. Pártay, and G. Csányi. Constant-pressure nested sampling with atomistic dynamics. *Physical Review E*, 96(4):043311, 2017.

[27] V. A. Mandelshtam and P. A. Frantsuzov. Multiple structural transformations in Lennard-Jones clusters: Generic versus size-specific behavior. *The Journal of Chemical Physics*, 124(20):204511, 2006.

[28] P. A. Frantsuzov and V. A. Mandelshtam. Size-temperature phase diagram for small Lennard-Jones clusters. *Physical Review E*, 72(3):037102, 2005.