Determining pressure-temperature phase diagrams of materials

Robert J. N. Baldock\textsuperscript{1,*}, Lívia B. Pártay\textsuperscript{2}, Albert P. Bartók\textsuperscript{3}, Michael C. Payne\textsuperscript{1}, and Gábor Csányi\textsuperscript{3}

\textsuperscript{1}Cavendish, \textsuperscript{2}University Chemical and \textsuperscript{3}Engineering Laboratories, University of Cambridge, Cambridge, United Kingdom

(Dated: April 28, 2016)

We extend the nested sampling algorithm to simulate materials under periodic boundary and constant pressure conditions, and show how it can be used to determine the complete equilibrium phase diagram, for a given potential energy function, efficiently and in a highly automated fashion. The only inputs required are the composition and the desired pressure and temperature ranges, in particular, solid-solid phase transitions are recovered without any \textit{a priori} knowledge about the structure of solid phases. We benchmark and showcase the algorithm on the periodic Lennard-Jones system, aluminium and NiTi.

I. INTRODUCTION

Phase diagrams of materials describe the regions of stability and equilibria of structurally distinct phases and are fundamental in both materials science and industry. In order to augment experiments, computer simulations and theoretical calculations are often used to provide reference data and describe phase transitions. A plethora of methods exist to determine individual phase boundaries, including Gibbs ensemble Monte Carlo [1], Gibbs–Duhem integration [2], thermodynamic integration and even direct molecular dynamics simulations of coexistence. Each of these algorithms requires the user to specify at least the identity and approximate location of the phase transition under investigation. Moreover, in the case of the solid phases, where much of the interest lies, advance knowledge of the crystal structure of each phase is required. Calculating an entire phase diagram by combining the results of such methods therefore demands a high degree of prior knowledge of the result. This in turn poses a barrier to the discovery of unexpected phases and phase transitions. Furthermore, such algorithms require specific expertise and separate setup for each type of phase transition.

In this paper we introduce a single algorithm, based on nested sampling (NS) [3, 4], that enables the efficient calculation of \textit{complete} pressure-temperature phase diagrams, including the solid region. This algorithm requires no prior knowledge of the phase diagram, and takes only the potential energy function together with the desired pressure and temperature ranges as inputs. Moreover, the direct output of the simulation is the potential energy function as an explicit function of its natural variables, so calculating thermodynamic observables, such as the heat capacity, is straightforward.

Nested sampling systematically explores the entire potential energy landscape, and in this way is related to parallel tempering (also known as replica exchange) [5, 6] and Wang-Landau sampling [7]. However, those algorithms encounter a particular convergence problem at first order phase transitions because the probability distributions (parametrised in terms of temperature in case of parallel tempering or energy in case of Wang-Landau) on the two sides of the phase transition have very little overlap [8]. This results in poor equilibration between the distributions on either side of the phase transition and large errors (both random and systematic) in the predicted locations of phase transitions.

The NS algorithm was designed to solve this problem. It constructs a sequence of decreasing potential energy levels, \{\textit{E}\textsubscript{i}\}, each of which bounds from above a volume of configuration space \chi\textsubscript{i}, with the property that \chi\textsubscript{i} is approximately a constant factor smaller than the volume, \chi\textsubscript{i−1}, corresponding to the level above. Each volume is sampled uniformly, and therefore each distribution will have an approximately constant fractional overlap with the one immediately before and after, ensuring fast convergence of the sampling and allowing an accurate evaluation of phase space integrals. In particular, the energy levels near the phase transition, where phase volumes change rapidly, will be very narrowly spaced. The sequence of energy levels comprise a discretisation of the cumulative density of states \chi(E), which allows the evaluation of the partition function at arbitrary temperatures,

\begin{align}
Z(N,V,\beta) &= \frac{1}{N!} \left(\frac{2\pi m}{\beta h^2}\right)^{3N/2} \int dE \chi'(E) e^{-\beta E} \\
&\approx Z_m(N,\beta) \sum_i (\chi_{i−1} − \chi_i) e^{-\beta E_i}
\end{align}

where \textit{N} is the number of particles of mass \textit{m}, \textit{V} is the volume, \beta is the inverse temperature, \textit{h} is Planck’s constant, the density of states \chi’ is the derivative of \chi, and we labelled the factor resulting from the momentum integral as \textit{Z}\textsubscript{m}. The total phase space volume is \chi\textsubscript{0} = V\textsubscript{N} corresponding to the ideal gas limit. Note that the sequence of energies \{\textit{E}\textsubscript{i}\} and configuration space volumes \{\chi\textsubscript{i}\} are independent of temperature, so the partition function can be evaluated \textit{a posteriori} at any temperature by changing \beta in (2).

The basic NS algorithm is as follows. We initialise by generating a pool of \textit{K} uniformly random configurations and iterate the following loop starting at \textit{i} = 1.

\textsuperscript{*} rjinbaldock@gmail.com
1. Record the energy of the sample with the highest energy as $E_i$, and use it as the new energy limit, $E_{\text{limit}} \leftarrow E_i$. The corresponding phase space volume is $\chi_i \approx \chi_0 [K/(K + 1)]^i$.

2. Remove the sample with energy $E_i$ from the pool and generate a new configuration uniformly random in the configuration space, subject to the constraint that its energy is less than $E_{\text{limit}}$. One way to do this is to clone a randomly chosen existing configuration and make it undergo a random walk of $L$ steps, subject only to the energy limit constraint.

3. Let $i \leftarrow i + 1$, and return to step 1.

At each iteration, the pool of $K$ samples are uniformly distributed in configuration space with energy $E < E_{\text{limit}}$. The finite sample size leads to a statistical error in log $\chi_i$, and also in the computed observables, that is asymptotically proportional to $1/\sqrt{K}$, so any desired accuracy can be achieved by increasing $K$. Note that for any given $K$, the sequence of energies and phase volumes converge exponentially fast (the number of iterations required to obtain results shown below never exceeded 2000 $\cdot K$), and increasing $K$ necessitates a new simulation from scratch.

Since its inception NS has been used successfully for Bayesian model selection in astrophysics [9], and also to investigate the potential energy landscapes of atomistic systems ranging from clusters to proteins [10–18].

The structure of this paper is as follows. In section II we modify the NS algorithm to enable its application at constant isotropic pressure with fully flexible periodic boundary conditions [19] where the periodic simulation cell is allowed to change shape. In sections III and IV we show that this development enables the determination of pressure-temperature phase diagrams of materials directly from the potential energy function without recourse to any other a priori knowledge. In particular, in section IV we calculate phase diagrams for aluminium and NiTi. Finally in section V we conclude this paper, discussing some consequences of the capability to calculate entire phase diagrams with a single method and in a highly automated fashion.

II. NESTED SAMPLING WITH FULLY FLEXIBLE PERIODIC BOUNDARY CONDITIONS AT CONSTANT PRESSURE

Nested sampling produces new samples by cloning an existing sample and then evolving the clone using a Markov chain Monte Carlo (MCMC) random walk [20]. Although one could work in the NVT ensemble and use equations 1 and 2, that would be very inefficient. MCMC simulations performed at fixed pressure require just a fraction of the computational expense as equivalent calculations performed at fixed volume. There are two reasons for this.

First, allowing the system to change volume by dilating or contracting expedites the cooperative freeing of jammed atoms. In contrast, at fixed volume, atoms that have become jammed are only freed by the coincidental movement of all atoms to separate them. Consequently MCMC simulations at fixed pressure explore configuration space far more rapidly than simulations at fixed volume.

The second reason arises from the thermodynamic behaviour of systems at a first order phase transition. At a phase transition under constant volume conditions the two phases coexist and an interface forms between them. Such interfaces are large on the atomic scale [21] and the behaviour of atoms at an interface is not representative of the behaviour of atoms in the equilibrium phases. As a result the interface introduces a systematic error that is only overcome by simulating very large numbers of atoms.

Such interfaces also occur under constant pressure conditions in the infinite system size limit. The contribution to the Gibbs Free Energy from an interface is proportional to $\gamma N^{2/3}$, where $\gamma$ is the interfacial tension. In contrast, the Gibbs Free Energies of each of the pure phases are extensive (proportional to $N$). Therefore the Gibbs Free Energy cost of the interface is negligible for thermodynamic systems. Conversely, for the relatively small system sizes amenable to density of states calculation methods such as nested sampling, the Gibbs Free Energy cost of the interface is appreciable, provided $\gamma$ is not close to zero. Consequently, at a constant pressure phase transition between phases with identical atomic compositions, configurations containing an interface have vanishing statistical weight in such simulations, and a discontinuous transition is observed from one equilibrium phase to the other. This enables the accurate simulation of phase transitions using much smaller numbers of atoms.

Using small numbers of atoms to simulate a phase transition naturally introduces new finite size errors. In particular, for a fixed number of atoms, it is not possible to represent all crystal structures in a simulation cell of fixed shape. This representational bias is removed by the use of fully flexible periodic boundary conditions [19], which allow the simulation cell to deform smoothly and thus take any shape. However, using fully flexible periodic boundary conditions allows the formation of very thin simulation cells containing unphysical quasi one and two dimensional configurations, characterised by interacting periodic images. In subsection II A we describe a rigorous solution to this new finite size problem. Later, in subsection II B we describe the calculation of the constant pressure partition function and heat capacity, both as explicit functions of temperature, using nested sampling.
A. Constraint on the simulation cell to exclude unphysical quasi one and two dimensional configurations

The partition function at fixed isotropic pressure \( p \) with fully flexible periodic boundary conditions [19] is

\[
\Delta(N, p, \beta) = Z_m \beta p \int dh_0 \delta (\det h_0 - 1) \times 
\int_0^\infty dVV^N \int_{(0,1)^{3N}} ds \, e^{-\beta H(s, h_0, V, p)},
\]

(3)

Here \( H(s, h_0, V, p) = E(s, h_0, V) + pV \), \( h \) is the \( 3 \times 3 \) matrix of lattice vectors relating the Cartesian positions of the atoms \( r \) to the fractional coordinates \( s \) via \( r = hs \), \( V = \det h \) is the volume, and \( h_0 = hV^{-1/3} \) is the image of the unit cell normalised to unit volume.

The partition function (3) corresponds to integration over all nine elements of the matrix \( h_0 \), and the \( \delta \)-function restricts the integration to matrices satisfying \( \det h_0 = 1 \).

This partition function is formally correct in the thermodynamic limit [19, 22]. However, finite systems in this description can adopt configurations for which the simulation cell becomes very thin. In this case, periodic boundary conditions give rise to a quasi one or two dimensional system. The prevalence of such configurations leads to a poor approximation of the three dimensional system. The prevalence of such configurations leads to a poor approximation of the three dimensional system due to excessively large finite size effects. We exclude such thin configurations by changing the limits for integration over elements of \( h_0 \), so that the perpendicular distances between opposite faces of the simulation cell \( h_0 \) are greater than some “minimum cell depth” value \( d_0 \).

The perpendicular distance between faces of the unit cell \( h \) made by lattice vectors \( h^{(i)} \) and \( h^{(j)} \) is given by

\[
d_{h^{(i) \perp}} = \frac{\det h}{|h^{(i)} \times h^{(j)}|}.
\]

(4)

The cell depth \( D(h_0) \), which measures how “thin” the cell has become, is defined as the minimum value of \( d_{h^{(i) \perp}} \), for the cell at normalised (unit) volume \( h_0 \).

\[
D(h_0) = \min_{i=1,2,3} \left( d_{h^{(i) \perp}} \right)
\]

(5)

Thus we integrate over elements of \( h_0 \) such that

\[
D(h_0) > d_0.
\]

(6)

The minimum cell depth \( d_0 \) is a real number on the interval \( [0, 1] \) where \( d_0 = 1 \) restricts the simulation cell to a cube. Smaller values of \( d_0 \) are accordingly less restrictive on the shape of the simulation cell, and \( d_0 = 0 \) corresponds to no restrictions on the simulation cell.

Incorporating this change of integration limits into the partition function (3) yields a new partition function

\[
\tilde{\Delta}(N, p, \beta, d_0) = Z_m \beta p \int_{D(h_0)>d_0} dh_0 \delta (\det h_0 - 1) \times
\int_0^\infty dVV^N \int_{(0,1)^{3N}} ds \, e^{-\beta H(s, h_0, V, p)}.
\]

(7)

In the thermodynamic limit (7) is equal to (3) up to a factor which depends only on \( d_0 \). The two partition functions are equal if and only if \( d_0 = 0 \).

In tests with 64 atoms we verified that the heat capacity curves were independent of \( d_0 \) at values of 0.65, 0.7 and 0.8, in Lennard-Jonesium and aluminium. The window of independence from \( d_0 \) grows wider as the number of particles is increased. For larger numbers of atoms, there are more ways to arrange those atoms into a given crystal structure, including in simulation cells that are closer to a cube. Similarly, unphysical correlations are introduced when the absolute number of particles between faces of the cell becomes too small, and therefore larger simulations can tolerate “thinner” simulation cells \( h_0 \).

The nickel-titanium calculations were performed with \( d_0 = 0.7 \).

B. Partition function and thermodynamic variables

The partition function we seek to calculate is given in equation (7). Above some sufficiently large volume \( V_0 \), we approximate the system as an ideal gas, neglecting interatomic interactions, which corresponds to the condition \( E(s, h_0, V) \ll pV \). In this approximation the volume integral in (7) is the sum of two parts

\[
\tilde{\Delta}(N, p, \beta, d_0) \approx Z_m \beta p \left[ \Delta_{NS}(N, p, \beta, V_0, d_0) + \int_{D(h_0)>d_0} \int_0^\infty dVV^N \int_{(0,1)^{3N}} ds \, e^{-\beta pV} \right]
\]

(8)

where

\[
\Delta_{NS}(N, p, \beta, V_0, d_0) = \int_{D(h_0)>d_0} \int_0^{V_0} dVV^N \int_{(0,1)^{3N}} ds \, e^{-\beta E(s, h_0, V_0, p)}
\]

(9)

We calculate \( \Delta_{NS} \) using nested sampling. Calculations are performed at fixed pressure to generate a sequence of enthalpies, \( H_i \), where \( H = E(s, V, h_0) + pV \). The NS approximation for \( \Delta_{NS} \), is

\[
\Delta_{NS}(N, p, \beta, V_0, d_0) \approx \sum_{i=1}^{\text{max}} (\chi_{i-1} - \chi_i) e^{-\beta H_i}
\]

(10)

where \( \chi_i \approx \chi_0 \left( \frac{K}{N+1} \right)^i \), \( \chi_0 = \frac{V_{N+1}}{N+1} \), and \( \Delta \chi_i \approx \chi_{i-1} - \chi_i \). We use single atom Monte Carlo (MC) moves in fractional coordinates with the amplitude updated every \( K \) iterations to maintain a good acceptance rate. Uniform sampling of lattice shape matrices \( h_0 \) subject to
equation (6) was achieved by independent shearing and stretching moves which do not change the volume. The ratios of atom, volume, shear and stretch moves were $N:10:1:1$. Further details of the MC moves and parallelisation scheme are given in the Supplemental Material (SM) [23].

We show in appendix A that volumes greater than $V_0$ make a negligible contribution to the partition function (8), provided $k_B T \ll pV_0$. In this case we have

\[
\tilde{\Delta}(N,p,\beta,d_0) \approx \beta p \left( \frac{2\pi m}{\beta h^2} \right)^{3N/2} N! \Delta_{NS}(N,p,\beta,V_0,d_0)
\]

where we have expanded $Z_m$. One can always assert the condition $k_B T \ll pV_0$, and in practice it is easy to find values of $V_0$ suitable for physically relevant conditions. We found $V_0 = 10^7 N \text{ Å}^3$ to be suitable for all conditions considered in this paper. From (11) we obtain the expected enthalpy

\[
\langle H \rangle = - \frac{\partial \log \tilde{\Delta}(N,p,\beta,d_0)}{\partial \beta}
\]

\[
= \left( \frac{3N}{2} - 1 \right) \frac{1}{\beta} + \langle H_{\text{configurations}} \rangle
\]

(12)

and the heat capacity at constant pressure

\[
C_p = -k_B \beta^2 \frac{\partial \langle H \rangle}{\partial \beta}
\]

\[
= \left( \frac{3N}{2} - 1 \right) k_B
\]

\[
+ k_B \beta^2 \left( \langle H^2_{\text{configurations}} \rangle - \langle H_{\text{configurations}} \rangle^2 \right)
\]

(13)

(14)

where

\[
\langle H_{\text{configurations}} \rangle \approx \sum_{i=1}^{i_{\text{max}}} \frac{\Delta \chi_i H_i e^{-\beta H_i}}{\sum_{i=1}^{i_{\text{max}}} \Delta \chi_i e^{-\beta H_i}}
\]

\[
\langle H^2_{\text{configurations}} \rangle \approx \sum_{i=1}^{i_{\text{max}}} \frac{\Delta \chi_i H_i^2 e^{-\beta H_i}}{\sum_{i=1}^{i_{\text{max}}} \Delta \chi_i e^{-\beta H_i}}
\]

(15)

This form (15) naturally does not depend on the contribution made by the low density configurations omitted from the NS calculation, or explicitly on the value of $d_0$. We used equations (15) when calculating the heat capacities presented in this paper.

### III. Calculating Phase Diagrams

In this section we describe a method for calculating the phase diagram of a material from the output of nested sampling. We then benchmark the performance of nested sampling on the periodic Lennard-Jones system, and find nested sampling to be orders of magnitude more efficient than Parallel Tempering (PT) for resolving the melting and evaporation transitions.

Given the partition function (11), phase transitions can be easily located by finding the peaks of response functions such as the heat capacity (14). By performing separate NS simulations at a number of pressures and combining the pressure and temperature values corresponding to the heat capacity peaks one can straightforwardly construct the entire phase diagram including all thermodynamically stable phases. This process is illustrated in Figure 1.

In Figure 2 we compare the performance of NS to that of PT for calculating the melting and evaporation transitions. NS provides a reasonable estimate of the melting and boiling points using only $\sim 10^8$ energy evaluations, while parallel tempering needs many orders of magnitude more computational effort than NS to find the evaporation transition and almost two orders of magnitude more computational effort to find the melting transition. (A similar increase in computational efficiency compared with parallel tempering was found for LJ clusters [10] and hard spheres [11, 27].)

Finally, in Figure 3 we show the phase diagram for 64 particles of Lennard-Jonesium as calculated using NS with $K = 640, L = 1.6 \times 10^5$. Comparison with the literature phase diagrams for $\sim 500$ particles confirms excellent agreement with the literature values for the evaporation transition [25] and also the solid-liquid and high pressure solid-vapour transitions [24]. Below the triple point, we observe slower convergence with respect to $L$ towards literature values of the sublimation transition [26]. We also find the beginning of the Widom-line: the shallow line of heat capacity maxima that extends into the supercritical region. The Widom-line and our method for estimating the critical point are described in the SM [23].
FIG. 2. Performance comparison of NS and PT. Sixty-four Lennard-Jones particles were simulated at a pressure of 0.027 (Lennard-Jones units). Both NS and PT simulations were initialised from the vapour phase. PT was performed using 128 equispaced temperature values in the range $[0.4, 1.4]$. The left panel shows the estimated transition temperatures as a function of computational cost while the right panel shows the mean enthalpy as a function of temperature corresponding to three selected values of the cost.

FIG. 3. Phase diagram for $N = 64$ Lennard-Jones particles as calculated using NS, with comparison to the literature ($N \approx 500$) phase diagram, as described in the text.

IV. RESULTS

A. Aluminium

In this section we apply the new algorithm to several empirical models of aluminium in order to demonstrate the capability of nested sampling to find solid-solid phase transitions without any prior knowledge of the crystal structures or even the existence of multiple stable phases. Furthermore, although the particular off-the-shelf models we use here do not reproduce the experimentally determined phase diagram of the material everywhere, the fact that nested sampling allows a direct calculation of the entire phase diagram means that in the future one could automate the optimisation of potentials to match the experimental phase diagram.

As one of the most commonly used metals, the thermodynamic properties of aluminium have been extensively studied. The melting line of aluminium has been measured up to 125 GPa [28–31], with good agreement between the different experimental techniques. Theoretical calculations have also been performed using embedded-atom type potentials [37–44] and ab initio methods [45–47], the latter providing melting temperatures up to 350 GPa [48]. At ambient conditions aluminium crystallises in the face-centred-cubic (fcc) structure, but a phase transition to the hexagonal-close-packed (hcp) structure at 217 GPa has been revealed by X-ray diffraction experiments [49] and the body-centred-cubic (bcc) phase has been also produced in laser-induced microexplosions [50]. The critical points of most metals are not amenable to conventional experimental study and thus estimation of their properties is usually based upon empirical relationships between the critical temperature and other measured thermodynamic properties. In the case of aluminium these result in predictions in a wide temperature and pressure range [32–35].

We chose four widely used models all based on the embedded-atom method (EAM): (1) the model developed by Liu et al. [43] (LEA-EAM), which is an improved version of the original potential of Ercolessi and Adams [42], (2) the model developed by Mishin et al. [44] using both experimental and ab initio data (Mishin-EAM), (3) the EAM of Mei and Davenport [40] (MD-EAM) and (4) the recently modified version of the MD-EAM, reparametrised by Jasper et al. to accurately reproduce the DFT energies for Al clusters and nanoparticles of various sizes (NPB-EAM) [51].

The phase diagrams for all four models based on NS simulations with 64 particles are shown in Figure 4. The resulting critical parameters vary over a wide range for the different models. Above the critical point we observe the Widom-line, indicated by those points not linked by a solid line. Heat capacity maxima corresponding to the Widom-line become broader away from the critical point, as indicated by the larger error bars. The Widom-line and our method for estimating the critical point are described in the SM [23].

The melting lines are in a good agreement with the available experimental data up to the pressure value $p \approx 25$ GPa. Above that pressure, the melting curves diverge from the experimental results, except for the MD-EAM potential, which reproduces the melting curve remarkably well.

At higher pressures small peaks appear on the heat capacity curves below the melting temperature for all models indicating solid-solid phase transitions (see appendix B). We post-processed the samples from the NS simulations. As expected, the fcc structure was found to be stable at low pressures in all four models. However, the models differ markedly in their predictions at high pressures. The only commonality between the predicted high pressure solid phase diagrams is that the maximum
predicted stable pressure for the fcc structure is far too low, both in comparison with experiment and density functional theory [49, 52, 53].

B. NiTi

Finally, in order to demonstrate that NS is applicable to more complex problems, we show results for a material of current scientific interest, the NiTi shape memory alloy [54, 55]. The shape memory effect relies on the structural phase transition from the high temperature austenitic phase to the low temperature martensitic phase [56]. Studying this transition is particularly challenging with traditional free energy methods because the austenitic phase does not correspond to a local minimum of the potential energy surface. Figure 5 shows the pressure-temperature-composition phase diagram corresponding to a recent EAM model [57, 58] as computed with NS. The phase transition temperatures are within 50 K of the experimental values and reproduce the trend with compositional change. We predict a decreasing transition temperature with increasing pressure. It is notable that this potential successfully reproduces the martensitic...
sition temperature, despite the fact that the minimum enthalpy structure for the potential is different to the structure observed both experimentally and in DFT: here the lowest enthalpy structure (which we label B19X) is orthorhombic (see the SM [23] for a description of the low enthalpy structures we identified). Thus it appears that the austenite-martensite transition temperature is not sensitive to the detailed geometry and ordering of the lowest enthalpy structures. Such empirical potentials can therefore be useful tools for studying this transition in the future.

V. CONCLUSION AND OUTLOOK

In summary, we have extended the nested sampling algorithm to allow simulations using fully flexible periodic boundary conditions at fixed pressure and demonstrated how it can be used to determine pressure-temperature-composition phase diagrams. In contrast to existing methods for comparing specific phases, NS explores the entire configuration space without requiring any prior knowledge about the structures of different solid phases with the only necessary input being the composition and the desired pressure and temperature ranges. This makes it the method of choice for exploring the pressure-temperature-composition space, which is the next unexplored realm naturally following much recent work in crystal structure exploration at zero temperature. Since the algorithm is run independently for different pressures and compositions, and also has excellent parallel scaling up to a number of processors equal to the number of simultaneous samples, it might even be possible to run it on ab initio models on exascale computers. Furthermore, we suggest NS is eminently suitable for validating materials models, and in the future could even play a role in the automatic optimisation of empirical models.

ACKNOWLEDGMENTS

RJNB acknowledges support from the EPSRC. LBP acknowledges support from St. Catharine’s College, Cambridge and to the Royal Society. APB acknowledges support from Magdalene College, Cambridge, the Leverhulme Trust and the Isaac Newton Trust. GC acknowledges EPSRC grant EP/J010847/1. Computer time was provided via the UKCP consortium funded by EPSRC grant ref. EP/K013564/1 and via the Darwin supercomputer in the University of Cambridge High Performance Computing Service funded under EPSRC grant EP/J017639/1. Data from this publication can be found at https://www.repository.cam.ac.uk/handle/1810/255091.

Appendix A: Ideal gas contribution to the partition function

In this appendix we show that the ideal gas contribution to the partition function (8) asymptotically approaches zero for any positive minimum cell depth \( d_0 \), in the limit \( k_B T/pV_0 \to 0 \). The ideal gas contribution to the partition function (8) is

\[
\int_{D(h_0)>d_0} dh_0 \delta (\det h_0 - 1) \int_{V_0}^{\infty} dVV^N \int_{(0,1)^3} ds e^{-\beta pV}
\]

We begin by noting that the exponential term does not depend on \( E(s,h_0,V) \), and therefore \( \int_{(0,1)^3} ds = 1 \). Thus we have

\[
\int_{D(h_0)>d_0} dh_0 \delta (\det h_0 - 1) \int_{V_0}^{\infty} dVV^N \int_{(0,1)^3} dse^{-\beta pV}
= \int_{D(h_0)>d_0} dh_0 \delta (\det h_0 - 1) \int_{V_0}^{\infty} dVV^N e^{-\beta pV}.
\]

The integral over volume \( V \) evaluates to

\[
\int_{V_0}^{\infty} dVV^N e^{-\beta pV} = \frac{1}{(\beta p)^{N+1}} \Gamma(N+1,\beta pV_0)
\]

where \( \Gamma(N+1,\beta pV_0) \) is the upper incomplete gamma function.

Finally we define the function \( A(d_0) \) to be equal to the integral over \( h_0 \)

\[
A(d_0) = \int_{D(h_0)>d_0} dh_0 \delta (\det h_0 - 1).
\]

The function \( A(d_0) \) is finite for any positive value of \( d_0 \), \( A(1) = 0 \) and \( A(d_0) \) diverges in the limit \( d_0 \to 0 \). In
the orthorhombic case, where all angles of the simulation cell are equal to $\frac{\pi}{2}$, $A(d_o) = \frac{d_o}{2} (\log d_o)^2$, with $A = 1$ at $d_o \approx 0.62$. However at any positive value of $d_o$ the contribution to the partition function (8) due to volumes greater than $V_0$ goes to zero in the limit $k_B T/pV_0 \to 0$ because $\Gamma(N + 1, \beta pV_0) \to 0$ in the same limit.

Appendix B: Identifying solid-solid phase transitions

The locations of phase transitions are determined solely by looking at the peaks in the heat capacity. Next, we inspect the system at temperatures either side of the phase transition. Specific phases can be identified in the following way. If no appropriate order parameter is to hand, then one picks a number of random configurations from the output of nested sampling, chosen according to their thermal weights $\Delta \chi e^{-\beta H}$, and inspects them by eye. If an appropriate order parameter is known, one can compute the free energy landscape for that order parameter. Here one proceeds by binning the weights $\Delta \chi e^{-\beta H}$ of all configurations, according to the order parameter, to create a partial sum $\Delta_j = \sum \Delta \chi e^{-\beta H}$ for each bin $j$. The free energy for each bin can then be computed as $F_j = -\frac{1}{\beta} \log(\Delta_j) + \log(\frac{\beta p}{N \beta}) + \frac{3N}{2} \log \left( \frac{2\pi m}{\beta n^2} \right)$. In fact, simply calculating the expected enthalpy at the phase transition, and then examining the order parameter values for output configurations around that enthalpy is often sufficient to identify the crystal structures.

An example of the latter approach is shown in Figure 6 for the Mishin-EAM potential, which compares the enthalpies and $Q_9$ bond order parameter values for nested sampling output configurations at three different pressures. At $p = 25.0$ GPa no phase transition occurs, and only fcc configurations are present. At $p = 34.9$ GPa a first order phase transition occurs at the average enthalpy marked by the vertical dashed line. At that enthalpy there is a clear transition between two basins, from a first basin that corresponds to the bcc structure, to a second that corresponds to the hcp structure. Finally, at $p = 37.5$ GPa no phase transition occurs and so there is no peak in the heat capacity. At this pressure the bcc structure is stable at all temperatures below the melting point. Nevertheless, the hcp structure is clearly visible as a metastable structure.

[23] See Supplemental Material at [URL will be inserted by publisher] for the detailed Monte Carlo scheme, parallelisation scheme, and additional results and details including treatment of the critical point, Widom-line, and density-temperature phase diagrams.
FIG. 6. Identification of solid phases from inspection of an order parameter: Mishin-EAM potential for aluminium. These plots compare the enthalpies and $Q_6$ bond order parameter values for nested sampling output configurations at three different pressures. Nested sampling proceeds towards lower enthalpies, thus from right to left in each plot. Horizontal dotted lines show the bond order parameters for the perfect fcc, bcc and hcp phases, and the vertical dashed line shows the expected enthalpy at the solid-solid phase transition, which was located by inspecting the heat capacity and observing a peak. These results are discussed in the text.
Determining pressure-temperature phase diagrams of materials: Supplemental Material


Cavendish, University Chemical and Engineering Laboratories, University of Cambridge, Cambridge, United Kingdom

(Dated: April 28, 2016)

DETAILED MONTE CARLO SCHEME

To understand the Monte Carlo scheme for NS in the fixed pressure ensemble with a fully flexible cell, it is helpful to first consider the Monte Carlo scheme for the fixed volume ensemble, with a fixed cell. In that case the partition function is given by

\[ Z(N, V, \beta) = \frac{1}{N!} \left( \frac{2\pi m}{\beta h^2} \right)^{3N/2} V^N \int ds \, e^{-\beta E(s, V)}. \] (S1)

As described in previous papers [S1–S3], each iteration begins with \( K \) samples distributed uniformly in \( s \), with \( E < E_{\text{limit}} \). That is to say, those \( K \) samples are drawn from the distribution

\[ \text{Prob}(s | E_{\text{limit}}) = \begin{cases} \frac{1}{\chi(E_{\text{limit}})}, & E(s, V) < E_{\text{limit}} \\ 0, & \text{Elsewhere} \end{cases} \] (S2)

For the fixed pressure ensemble, we must explore volume and cell matrices \( h_0 \), as well as the fractional coordinates \( s \). In particular, we calculate the integral (9) with NS.

\[ \Delta_{\text{NS}}(N, p, \beta, V_0, d_0) = \int_{D(h_0) > d_0} dh_0 \delta(\det h_0 - 1) \times \int_0^{V_0} dV V^N \int_{(0,1)^3} ds \, e^{-\beta [E(s, h_0, V) + pV]} \] (9, revisited)

In (9) the volume has a weight \( V^N \) with \( V < V_0 \), and \( h_0 \) is constrained to a surface of determinant 1, with \( D(h_0) > d_0 \). Therefore we seek to generate new samples according to the distribution

\[ \text{Prob}(s, h_0, V | H_{\text{limit}}) = \begin{cases} \frac{V^N \delta(\det h_0 - 1)}{\chi(H_{\text{limit}})}, & H(s, h_0, V) < H_{\text{limit}}, \ V < V_0, \\ 0, & D(h_0) > d_0 \end{cases} \] (S3)

As described in the paper, this can be achieved by Markov chain Monte Carlo (MCMC) sampling. MCMC walks were random sequences of the following Monte Carlo (MC) moves, which were chosen according to the probabilities (ratios) given in the paper, unless specifically stated in this Supplemental Material.

Atom moves: Single atom MC moves are performed as follows.

1. Select a single atom at random. This atom has fractional coordinates \( s_i \).
2. Displace that atom’s fractional coordinates by a random vector
   \[ s_i \rightarrow s_i + L_s \Delta_s \] (S4)
   Here \( \Delta_s \) is a 3-tuple with elements drawn independently from a uniform distribution on the interval \([-1,1]\), and \( L_s \) is a constant that controls the step size. The size of \( L_s \) was updated periodically during the NS calculation, as described in the next subsection, “Updating MC step lengths”.
3. Calculate the enthalpy \( H_{\text{trial}} = E(s, h_0, V) + pV \) for the system after this atom displacement.
4. Accept the displacement if \( H_{\text{trial}} < H_{\text{limit}} \). Otherwise, the old configuration is kept.

Volume moves: Volume MC steps are as follows.
1. Propose a volume move from $V$ to $V'$ where

$$V' = V + L_V \Delta_V$$

(S5)

Here $\Delta_V$ is a random number drawn from the uniform distribution on the interval $[-1, 1]$, and $L_V$ is a constant that controls the step size. The size of $L_V$ was updated periodically during the NS calculation, as described in the next subsection, “Updating MC step lengths”.

2. The displacement is accepted with probability $\min\left(1, \left[\frac{V'}{V}\right]^N\right)$. If the displacement is accepted, continue to step 3. Otherwise, the old configuration is kept and the move ends.

3. If the displacement was accepted in step 2, calculate the enthalpy $H_{\text{trial}} = E_0(s, h_0, V') + pV'$ for the system after this volume displacement.

4. If $H_{\text{trial}} < H_{\text{limit}}$ the displacement is accepted. Otherwise, the old configuration is kept.

**Lattice shear moves:** The unit simulation cell $h_0$ is sheared, subject to constraints on both the cell depth and the enthalpy. The algorithm is given in Algorithm 1. The size of shear steps is controlled by the parameter $\text{step}_l\text{sh}$. This parameter was updated periodically during the NS calculation, as described in the next subsection, “Updating MC step lengths”.

**Lattice stretch moves:** The unit simulation cell $h_0$ is stretched at fixed unit volume, subject to constraints on both the cell depth and the enthalpy. The algorithm is given in Algorithm 2. The size of stretch steps is controlled by the parameter $\text{step}_l\text{st}$. This parameter was updated periodically during the NS calculation, as described in the next subsection, “Updating MC step lengths”.

**(Binary) atom swaps:** For the NiTi alloy, we included binary (Ni–Ti) atom swaps as an additional MC proposal. The algorithm for this MC proposal is as follows:

1. Choose a random Ni atom, and a random Ti atom.
2. Swap the fractional coordinates of those two atoms.
3. Calculate the enthalpy $H_{\text{trial}} = E_0(s, h_0, V) + pV$ for the system after the swap.
4. Accept the swap if $H_{\text{trial}} < H_{\text{limit}}$. Otherwise, the old configuration is kept.

**Updating MC step lengths**

The values of the step size parameters, $L_s$, $L_V$, $\text{step}_l\text{sh}$ and $\text{step}_l\text{st}$, were updated every $\frac{K}{2}$ iterations (recall that $K$ is the number of samples employed in the NS calculation). This corresponds to an expected reduction in the enclosed configuration space volume $\chi$ to a factor of $e^{-\frac{1}{2}}$. Each step size parameter was updated by performing a short MCMC exploration using only the corresponding type of MC proposal. The step size parameters were updated to maintain an acceptance rate in the interval $(0.2, 0.3)$, and as close to 0.25 as possible, for each MC proposal type.

Since the calculation was performed in parallel, when collecting data about the acceptance rate for each MC proposal type, each processor ran MCMC trajectories with separate, random configurations from the current sample set. In general, this avoids setting step lengths appropriate to unrepresentative regions of the bounded uniform distribution.

**PARALLELISATION SCHEME**

In order to be able to benefit from supercomputing facilities efficiently, the NS algorithm has to be parallelised. Fortunately, possible parallelisation schemes come quite naturally from the method, and here we describe a scheme for parallelising within each iteration [S4]. While in the serial algorithm, the new clone is decorrelated using a trajectory of walk length $L$ in each iteration, in our scheme parallelised over $N_{\text{proc}}$ processors, the new clone is decorrelated using a trajectory of walk length $L/N_{\text{proc}}$, along with $(N_{\text{proc}} - 1)$ randomly chosen configurations which are also (independently) propagated through $L/N_{\text{proc}}$ steps. Thus, rather than propagating a single configuration for $L$ steps, we propagate $N_{\text{proc}}$ configurations for $L/N_{\text{proc}}$ steps.
subroutine lattice_shear_mc
! lattice shear MC step
! The input unit cell matrix is h0_in(3,3)

h0_trial = h0_in           ! copy h0_in into working array

do repeat = 1, 3
    ! repeat three times

h0_save = h0_trial           ! save h0 in case of rejection

chng_vec = ranf()*3 + 1      ! randomly selected lattice vector: integer: 1, 2, or 3

basis = orthonormal_basis(h0_in,chng_vec)
! basis(3,2) is a pair of orthonormal vectors in the plane
! defined by the two edge vectors in h0_in other than chng_vec

pair: do
    x = 2.0*(ranf() - 0.5)   ! generate (x,y) pair from
    y = 2.0*(ranf() - 0.5)   ! inside unit circle
    if (x**2 + y**2 < 1.0) exit pair
end do pair

x = x*step_l_sh            ! x,y updated to uniform pair in
y = y*step_l_sh            ! a circle of radius step_l_sh

h0_trial(:,chng_vec) = h0_trial(:,chng_vec)
    + x*basis(:,1) + y*basis(:,2)
! change to cell vector chng_vec is circularly symmetric
! and in the plane defined by the other two edge vectors

if ( ch(h0_trial) < mch ) h0_trial = h0_save
! trial rejected due to minimum cell depth criterion.
! cell depth of proposed cell is smaller than minimum
! allowed value: D(h0) < d0

end do

H_trial = H(h0_trial)       ! H = configurational enthalpy
if ( H_trial <= H_lim ) then
    h0_out = h0_trial         ! accept trial
else
    h0_out = h0_in            ! trial rejected
    ! h0 unchanged: a null step
endif
end subroutine

Algorithm 1. Lattice shear MC step. Samples the surface det(h0) = 1 subject to the enthalpy and minimum cell depth constraints. ranf() is a random number uniform in [0,1].

FINITE SIZE EFFECT

In order to estimate the effect of using only 64 particles in our simulations we performed tests with 32 and 128 particles as well for both the periodic Lennard-Jones and the LEA-EAM aluminium potential models. The heat capacities of these are compared in Figure S1. The Lennard-Jones test was performed at $p = 0.027\epsilon/\sigma^3$, below the critical point, hence both the evaporation and melting peaks are present. The aluminium test was performed at $p = 11.6$ GPa, above the critical point, thus only the melting peak is present. The figure shows that increasing the number of atoms, the heat capacity peaks become narrower as expected, and the melting and evaporation transitions
Algorithm 2. Lattice stretch MC step. Samples the surface \( \det(h_0) = 1 \) subject to the enthalpy and minimum cell depth constraints. \( \text{ranf}() \) is a random number uniform in \([0,1]\).

shift to lower and higher temperatures respectively. Of course, it would be possible to perform a finite size scaling analysis from the results of multiple Nested Sampling calculations, performed at different system sizes. However, we felt that the phase diagrams of the various aluminium models differed from each other to the extent that it was sufficient to give a rough indication of error based on the full-width-at-half-maximum of the peaks.

LENNARD-JONES POTENTIALS

The benchmark comparison to parallel tempering was performed using the truncated and shifted Lennard-Jones potential [S5]. Other Lennard-Jones calculations were performed using the truncated but not shifted potential with the usual mean-field correction [S5] beyond the cut-off radius. All Lennard-Jones calculations employed a cut-off radius \( r_c = 3\sigma \).
FIG. S1. Magnitude of the finite size effect in case of the Lennard-Jones potential (left panel) run with parameters \( K = 672 \) and \( L = 15120 \), and LEA-EAM aluminium potential (right panel) with run parameters \( K = 1608 \) and \( L = 10080 \), as demonstrated by the heat capacity curves of systems of 32, 64 and 128 particles.

ALUMINIUM

Simulation details

The simulation cell contained 64 aluminium atoms. The same ratio of the different types of MC moves was used at every pressure. Sampling is robust with respect to the ratio of MC moves of each type, and for these aluminium calculations the ratio of MC moves (atom:volume:shear:stretch) was (16:1:1:1), rather than (N:10:1:1). The minimum allowed cell depth, \( d_0 \), was set to 0.65. Spline parameters for the LEA-EAM and Mishin-EAM potentials are only provided for configurations where no two atoms are closer than a certain, small radius. For separations smaller than this radius we made linear extrapolations to the splines. The ultra-close atomic proximity region is only relevant to the ultra-high temperature gas phase, and does not contribute to the thermodynamics of aluminium in the temperature and pressure ranges considered in this paper. Therefore the precise form of this extrapolation does not affect any of the results we show.

Heat capacity curves

Heat capacity curves calculated with Nested Sampling are shown in Figures S2, S3, S4 and S5.

Critical point and the Widom-line

To locate the critical point, we draw on the results of Bruce and Wilding [S6]. Specifically we begin from their analysis of the probability distribution for density for finite systems, in the pressure-temperature region that would contain the critical point in an infinite system. For a finite system at and below the critical point the density distribution for temperatures corresponding to maxima of the heat capacity appears as a bimodal distribution. Above the critical point (along the Widom-line), the distribution transitions quickly to a unimodal distribution. In the context of a phase diagram, our location of the critical point appears quite precise, but in fact our sampling of the pressure axis is relatively sparse compared to the width of the scaling region around the critical point, since we are inspecting the system across pressures spanning many orders of magnitude. Therefore, between adjacent pressure values the density distribution transitions rapidly between being bimodal and unimodal. We chose the lowest pressure for which the distribution is clearly unimodal as an upper bound for the critical pressure. The corresponding upper bound for the critical temperature was taken to be the temperature at which the heat capacity, \( C_p \), is its local maximum at that pressure.

According to textbook definitions, beyond the critical point a single fluid phase is defined. However, response functions such as the heat capacity, thermal expansion coefficient and compressibility continue to exhibit maxima into the supercritical region. Lines of these maxima, which spread out from the critical point are called the Widom-lines. If one moves away from the critical point, the different Widom-lines rapidly diverge from one another, and the maxima themselves become quickly smeared and disappear, in case of the heat capacity at \( T \approx 2.5T_c \) and \( p \approx 10p_c \) [S7].
FIG. S2. Aluminium heat capacity curves at lower pressures up to the end of the Widom-line for the different studied potential models. The arrows indicate the peak of the $c_p$ curves closest to the critical point, and in order to show the broadened peaks above that, the curves in the supercritical region are shown on a different scale (shown on the right hand side). The size of the live set and the total number of MC steps during one iteration was $K=804$ and $L=3120$, respectively.

FIG. S3. Aluminium heat capacity curves above the critical point for the LEA-EAM potential. The insets show the solid-solid phase transition peaks in a larger scale. The size of the live set was $K = 3216$ and the total number of MC steps during one iteration was $L = 15120$, or if otherwise, it is indicated in the legend.

The same tendency can be observed in our Nested Sampling results, see Fig. S2.

By calculating the expected density as a function of temperature, the equation-of-state can be calculated using Nested Sampling. These are plotted for the low pressure region together with Gibbs ensemble Monte Carlo simulation results from Ref [S8] in Figure S6. The critical parameters for the different aluminium potentials determined by Nested Sampling are summarised in Table S1.
FIG. S4. Aluminium heat capacity curves above the critical point for the NPB-EAM potential. The inset shows the solid-solid phase transition peaks in a larger scale. The size of the live set was \( K = 3216 \) and the total number of MC steps during one iteration was \( L = 15120 \).

FIG. S5. Aluminium heat capacity curves above the critical point for the MD-EAM potential. The size of the live set was \( K = 3216 \) and the total number of MC steps during one iteration was \( L = 15120 \).

**NITI**

Here we discuss the lowest enthalpy structures we identified for the EAM potential used to describe NiTi [S9]. The enthalpic and energetic ordering of these structures is the same for all pressures considered in this paper. Therefore we shall restrict our discussion to their energy values.

The EAM potential gives rise to a number of low energy structures that show some similarity to the experimentally observed low temperature structure (B19') and to other structures found using DFT (namely B19 and BCO). The lowest energy structure of the EAM potential is orthorhombic, but with significant displacement of the atoms from their high temperature average B2 positions. Table S2 below gives the structural parameters of the three structures with the lowest energies. In addition, we found that a multitude of local minima exist between just 10 and 15 meV/atom above the lowest energy structure. The symmetries of these configurations were identified using the findsym package [S10].

<table>
<thead>
<tr>
<th>Potential</th>
<th>( T_c ) (K)</th>
<th>( p_c ) (GPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEA-EAM</td>
<td>6810 ± 170</td>
<td>0.262</td>
</tr>
<tr>
<td>Mishin-EAM</td>
<td>5800 ± 140</td>
<td>0.054</td>
</tr>
<tr>
<td>NPB-EAM</td>
<td>6100 ± 150 (6299 ± 48)</td>
<td>0.083 (0.0896 ± 0.0019)</td>
</tr>
<tr>
<td>MD-EAM</td>
<td>3340 ± 140 (3381 ± 13)</td>
<td>0.042 (0.044 ± 0.002)</td>
</tr>
</tbody>
</table>
FIG. S6. Density-temperature phase diagram of aluminium with the NPB-EAM potential. Red lines show the Nested Sampling results at different pressures: 0.000402 GPa, 0.0012 GPa, 0.0208 GPa, 0.0337 GPa, 0.0503 GPa, 0.0679 GPa, 0.0830 GPa, 0.100 GPa, 0.115 GPa, 0.150 GPa, 0.230 GPa, 0.399 GPa, 0.807 GPa, 1.0 GPa, 2.5 GPa. The curve corresponding to the critical pressure is shown by a dashed line. Black and open circles show the liquid and vapour densities respectively, from [S8], with a black star showing their estimate of the critical point. The inset shows the minimum value of the derivative of the curves obtained by Nested Sampling, with the arrow indicating the pressure value where the plateau between vapour and liquid phases diminish at the critical point.

TABLE S2. Low energy structures of NiTi according to the EAM potential. Lattice vectors are in Å, energies are per atom and are relative to the B2 phase.

<table>
<thead>
<tr>
<th>Symmetry</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>β</th>
<th>Wyckoff positions</th>
<th>Energy (meV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pm3m (B2)</td>
<td>3.0098</td>
<td>3.0098</td>
<td>3.0098</td>
<td>90</td>
<td>Ni: a, Ti: b</td>
<td>0</td>
</tr>
<tr>
<td>Pmm2 (B19X)</td>
<td>4.02186</td>
<td>4.40909</td>
<td>3.01484</td>
<td>90</td>
<td>Ni: b (z=0.04073), c (z=0.04305) Ti: a (z=0.4434), d (z=0.4458)</td>
<td>-53.8</td>
</tr>
<tr>
<td>C2/m (B19)</td>
<td>4.46697</td>
<td>4.02169</td>
<td>3.00532</td>
<td>98.1</td>
<td>Ni: a, Ti: d</td>
<td>-51.1</td>
</tr>
<tr>
<td>P21/m (BCO)</td>
<td>2.99360</td>
<td>4.00083</td>
<td>4.88392</td>
<td>107.8</td>
<td>Ni: e (x=0.08755, z=0.3249) Ti: e (x=0.3583, z=0.2167)</td>
<td>-46.8</td>
</tr>
</tbody>
</table>