A fast anharmonic free energy method with an application to vacancies in ZrC

Thomas A. Mellan,[1, *] Andrew I. Duff,[2] Blazej Grabowski,[3] and Michael W. Finnis[1]

[1]*Thomas Young Centre, Department of Physics and Department of Materials,*
Imperial College London, Exhibition Road, London SW7 2AZ, United Kingdom
[2]*STFC Hartree Centre, Scitech Daresbury, Warrington WA4 4AD, United Kingdom*
[3]*Max-Planck-Institut für Eisenforschung GmbH, D-40237 Düsseldorf, Germany*

We propose an approach to calculate the anharmonic part of the volumetric-strain and temperature dependent free energy of a crystal. The method strikes an effective balance between accuracy and computational efficiency, showing a ×10 speed-up on comparable free energy approaches at the level of density functional theory, with average errors less than 1 meV/atom. As a demonstration we make new predictions on the thermodynamics of substoichiometric ZrC_x, including vacancy concentration and heat capacity.

I. INTRODUCTION

Thanks to recent advances in computational thermodynamics, the thermal properties of metals such as aluminium and gold have been investigated up to the melting point, using thermodynamic integration (TI) with Langevin dynamics[1–3] based on density functional theory (DFT). A two-step TI approach increases computational efficiency further, making predictions possible for more complex materials at the DFT level of theory. Examples so far include the ultra-high-temperature ceramics ZrC ($T_m = 3700\,\mathrm{K}$) and HfC ($T_m = 4160\,\mathrm{K}$),[4,5] and recent attempts at tackling the emerging class of multicomponent systems.[6,7] Such calculations are not yet routine, but the course of our research and the recent methodological developments of others[8–10] in this field is in that direction. Here we present some new developments that are a step towards the goal of routinely computing accurate free energies for hard matter systems, including binaries, ternaries, and high-entropy alloys, across the range of temperatures, pressures and chemical potentials, up to and eventually beyond the melting point.

In this work we compute the concentration of vacancies in ZrC_x and associated ambient pressure thermodynamics for small deviations from stoichiometry. The ZrC_x free energy and derivatives are analyzed in terms of the contributions[11]

$$F = E_0 + F_{el} + F_{qh} + F_{ah} + F_{el\text{-}vib} + F_{config}, \quad (1)$$

in which E_0 is the DFT energy of a static lattice at $T = 0\,\mathrm{K}$, F_{el} is the Helmholtz free energy contribution from the thermal excitations of electrons, F_{qh} is the quasiharmonic vibrational contribution, F_{ah} is the anharmonic vibrational contribution, $F_{el\text{-}vib}$ is the electron-vibration contribution, and F_{config} is the contribution of configurational entropy due to the number of distinct point-defect distributions. For each of the five temperature-dependent terms we have calculated the dependence on the independent variables volume and temperature up to the melting point. Considerable attention in this paper is given to the

method we use to compute the challenging anharmonic term, F_{ah}. The method described achieves effective DFT accuracy in F_{ah} (1 meV per bulk atom) with only an order of magnitude greater computational cost than ordinary quasiharmonic free energy calculations.

This paper is set out as follows. Sec. II A gives the context of our approach, theoretical details are in Sec. II B, a description of the modified embedded atom method (MEAM) potential fitting, which reduces overall the number of expensive DFT calculations, in Sec. II C, thermodynamic integration in Sec. II D, and DFT technical details in Sec. II E. Benchmarking is described in terms of accuracy and precision in Sec. III A and computational efficiency in Sec. III B. Application to ZrC_x provides insight into the nature of anharmonicity in substoichiometric binary crystals in Sec. IV A, prediction of vacancy concentration in Sec. IV B, and analysis of ZrC_x heat capacity in Sec. IV C.

II. METHODS

A. Background

There are a number of approaches in the literature to calculate the anharmonic vibrational properties of crystals,[1–4,12–22] including thermodynamic integration[23] (TI), which is the method used in this work. In TI the anharmonic part of the full Hamiltonian, $E - E_{qh}$, is switched on with the parameter $\lambda \in [0, 1]$, in this instance linearly as $E_{mix}(\lambda) = E_{qh} + \lambda(E - E_{qh})$. Classical averages of $\partial_\lambda E_{mix}(\lambda)$ are obtained stochastically from molecular dynamics (MD), and numerically integrated along the coupling path to give the free energy due to E_{ah}:

$$F_{ah} = \int_0^1 d\lambda \, \langle \partial_\lambda E_{mix}(\lambda) \rangle_\lambda . \quad (2)$$

Note, $\partial_\lambda E_{\mathrm{mix}}(\lambda) = E(\mathbf{R}, V) - E_{\mathrm{qh}}(\mathbf{R}, V)$, where $E(\mathbf{R}, V)$ is the full potential energy surface and $E_{\mathrm{qh}}(\mathbf{R}, V)$ the volume-dependent harmonic potential energy surface in Born-Oppenheimer nuclear coordinates $\mathbf{R}$.

The partitioning of the vibrational free energy $F_{\mathrm{vib}} = F_{\mathrm{qh}} + F_{\mathrm{ah}}$ divides the problem conveniently into a simple quantum mechanical part, in which the vibrations are quantised as phonons, and an anharmonic part in which the vibrations are treated classically. Thus F_{vib} has the appropriate low-temperature quantum statistics. The anharmonicity is treated classically but also non-perturbatively, which is important at high temperatures as the melting point is approached. In order to evaluate the anharmonic term, the expectation values $\langle \partial_\lambda E_{\mathrm{mix}}(\lambda) \rangle_\lambda$ require between $10^3 \ldots 10^7$ configurations for a typical λ-ensemble at a typical supercell size. To produce a free energy surface $F_{\mathrm{ah}}(V, T)$ one must sample ensembles across dimensions of strain (here volume), temperature and coupling parameter, $N_\lambda \times N_V \times N_T \approx 10^2 \ldots 10^3$. Thus the ball-park number of total energy calculations, between $10^5 \ldots 10^{10}$ configurations, is prohibitive at the highly-converged DFT level of accuracy required.

In one approach to reduce computational complexity, F_{ah} is obtained by cumulating a sequence of thermodynamic integrations. In an implementation of this approach referred to as TU-TILD,[4] which is expressed by Eqn. (3), much of F_{ah} is captured using an inexpensive MEAM potential. This results in faster convergence of the expensive TI from MEAM to DFT, expressed in the last term of Eqn. (3).

$$
\begin{aligned}
F_{\mathrm{ah}}^{\mathrm{TUTILD}} &= \int_0^1 \mathrm{d}\lambda \langle E^{\mathrm{DFT}}(\mathbf{R}, V) - E_{\mathrm{qh}}^{\mathrm{DFT}}(\mathbf{R}, V) \rangle_\lambda \\
&= \int_0^1 \mathrm{d}\lambda \langle E^{\mathrm{MEAM}}(\mathbf{R}, V) - E_{\mathrm{qh}}^{\mathrm{DFT}}(\mathbf{R}, V) \rangle_\lambda \\
&+ \int_0^1 \mathrm{d}\lambda \langle E^{\mathrm{DFT}}(\mathbf{R}, V) - E^{\mathrm{MEAM}}(\mathbf{R}, V) \rangle_\lambda .
\end{aligned}
\tag{3}
$$

In practice to save computation time, the DFT MD calculations in a TU-TILD procedure were usually performed with a low-converged expansion of the wavefunctions, using a reduced number of plane-waves, and fewer k-points than required for maximum accuracy. The maximum accuracy was then obtained by up-sampling, as in the original UP-TILD method[1]. The methodology we introduce below, inspired by these approaches, was devised in order to make significant further savings in computation time without sacrificing accuracy.

B. MEAM thermodynamic integration approach

The approach we propose in this work can be summarized by

$$
F_{\mathrm{ah}} = \int_0^1 d\lambda \left\langle E^{\mathrm{MEAM}}(\mathbf{R}, V) - E_{\mathrm{qh}}^{\mathrm{DFT}}(\mathbf{R}, V) \right\rangle_\lambda . \tag{4}
$$

Our approach calculates the anharmonic free energy of a MEAM crystal referenced to a harmonic DFT crystal, which is formally the first stage in Eqn. (3). In the present method the quasiharmonic Helmholtz free energy at each volume is still explicitly represented by the volume-dependent dynamical matrix calculated with DFT, which captures much of the thermal expansion, but the anharmonic terms are now entirely described by the MEAM potential. The success of the method depends on being able to generate a MEAM potential of sufficient accuracy to replace the anharmonic DFT contribution. It is by no means obvious *a priori* that this is possible, or if it is, that the process of generating the potential is not too expensive to warrant the effort.

From the potential terms in Eqn. (4) it is clear that F_{ah} can be evaluated by this method to a high level of precision using modest computational resources, but the MEAM TI approximation introduces systematic potential errors with respect to DFT TI. Accuracy must be carefully controlled by generating custom MEAM potentials from high-quality DFT MD. Generating the training and validation data is time consuming, so fitting the potentials becomes the primary computational cost in predicting F_{ah} in our TI approach. These costs incurred before doing the MEAM TI will be shown in Sec. III A to be comfortably small enough. Details of potential fitting and error control are presented in the following section.

C. Potential fitting

Interatomic potentials have been fitted using the reference-free modified embedded atom method (RF-MEAM).[24] The MEAM potentials fitted in this work lack transferability and are specialized to perform for the intended application. For instance, to model $F_{\mathrm{ah}}(V, T)$ we fit a separate potential at each volume considered, which minimizes the possibility of strain-dependent errors. Obtaining the correct implicit volume dependence of a potential is important, as explicit anharmonic effects depend sensitively on the degree of lattice expansion, as demonstrated in Sec. IV A.

Potentials for $\mathrm{Zr}_{32}\mathrm{C}_{32}$ and $\mathrm{Zr}_{32}\mathrm{C}_{31}$ have been fitted at the cell lattice parameters $a = \{4.685, 4.730, 4.759, 4.801, 4.850\}$ Å. In crystals of lower than cubic symmetry, thermal expansion may involve other modes of strain, but in our case only volumetric strain need be considered. At each volume the potential is fitted to configurations sampled from DFT MD runs between $T = 200\,\mathrm{K}$ and $T = 3800\,\mathrm{K}$. The fitting set for each volume comprises 10^3 $\mathrm{Zr}_{32}\mathrm{C}_{31}$ configurations, which supply energies and forces to
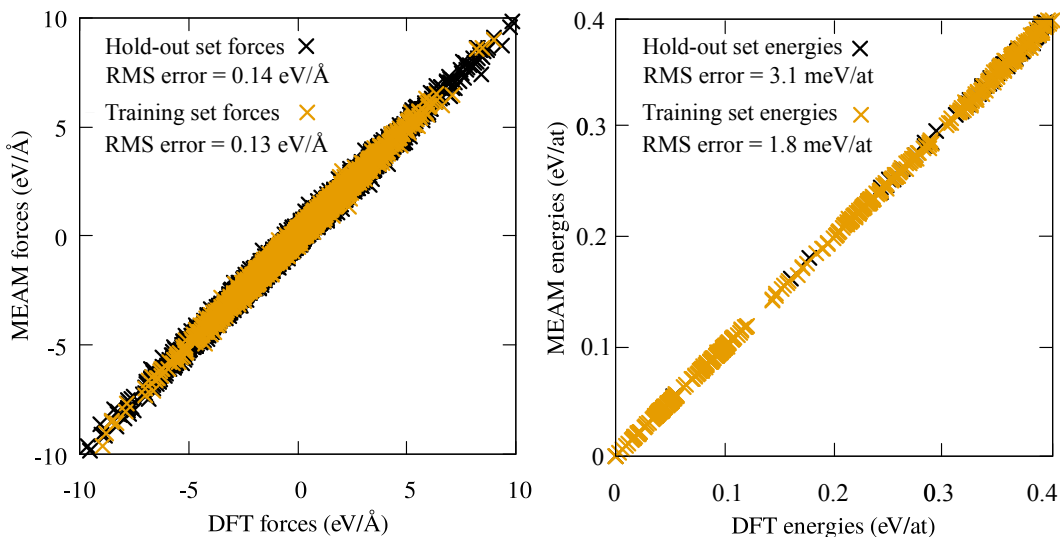
Figure 1. Quality of interatomic potential: MEAM versus DFT forces and energies shown for training and hold-out $Zr_{32}C_{31}$ data sets.

the objective function, for minimizing force residuals and energy-residual variances. Fitting to the forces on each atom allows the potential to be specified by fewer distinct configurations than fitting to energies, as $N_{forces} = 3N_{atom}N_{energies}$. To generate the fitting data points efficiently, low-quality DFT can be up-sampled to high-quality DFT to produce high-quality target forces and total energies.

The interatomic potentials are generated using a genetic-algorithm conjugate-gradient fitting procedure implemented and publicly available in the MEAMFIT2 code.[24–26] The code fits an RF-MEAM potential that permits locally positive and negative density terms in order to increase variational freedom, subject to a net positive background. The fitted potential has 3 embedding and 3 pairwise terms, within a radial cutoff of 4.8 Å, which includes interactions up to third nearest-neighbor. These potential parameters provide a satisfactory compromise between accuracy and complexity, in terms of minimizing residual variances on hold-out data using the fewest degrees of freedom (78 parameters for the 3-3 potential). The quality of fits for energy and forces is presented in Fig. 1 for $Zr_{32}C_{31}$.

In this RF-MEAM application a different potential is fitted at each volume but we require a potential to be transferable across composition, i.e. we want the same potential to describe both $Zr_{32}C_{31}$ and $Zr_{32}C_{32}$, for a given volume at any temperature. This transferability ensures a systematic error cancellation in F_{ah} for $Zr_{32}C_{31}$ and $Zr_{32}C_{32}$ that helps in calculating accurate fully anharmonic vacancy formation energies.
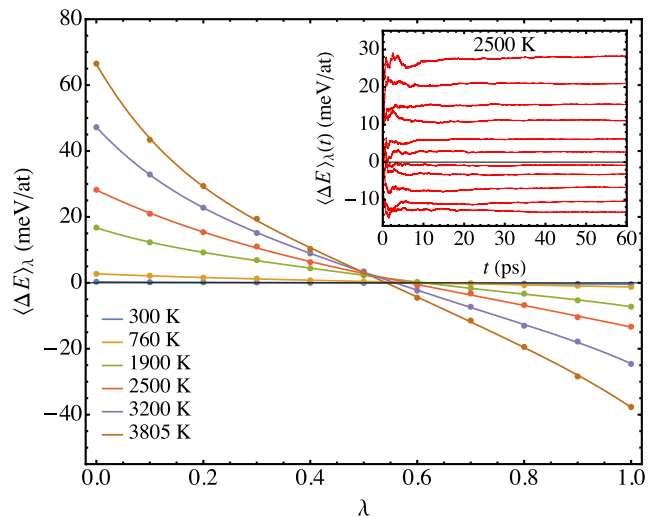
D.   Thermodynamic integration

Figure 2. Thermodynamic integrand $\langle\Delta E\rangle_\lambda (\lambda)$ for $Zr_{32}C_{31}$ at $a = 4.801$ Å. *Inset*: Convergence of $\langle\Delta E\rangle_\lambda (t)$ with MD time-step, at ensembles $\lambda_i = i/10$.

F_{ah} is estimated by computing $\langle\Delta E\rangle_\lambda$, with $\Delta E = E^{MEAM} - E_{qh}^{DFT}$, for a series of 11 ensembles at equal increments of λ, $\lambda_i = i/10$. In Fig. 2 we show the dependence of $\langle\Delta E\rangle_\lambda$ on λ across a series of temperatures. We see that at each temperature, $\partial_\lambda\langle\Delta E\rangle_\lambda \leq 0$, a necessary condition that is easy to prove, as in a derivation of the Gibbs-Bogoliubov inequality. In Fig. 2 inset the convergence of $\langle\Delta E\rangle_\lambda$ is shown for the first 60,000 time steps of a simulation. Expectation values are generated using Langevin MD, with a one femtosecond time-step,[4]

and a friction parameter of $\gamma = 0.05\,\text{fs}^{-1}$ for $\text{Zr}_{32}\text{C}_{31}$ and $\gamma = 0.01\,\text{fs}^{-1}$ for $\text{Zr}_{32}\text{C}_{32}$. At each pair, $\{V_i\,T_i\}$, the expectation value $\langle \Delta E \rangle_\lambda$ is fitted in λ by least squares to the truncated power series

$$\langle \Delta E \rangle_\lambda (\lambda) = \sum_{i=0}^{i=5} a_i \lambda^i, \ \lambda \in [0,\, 1], \tag{5}$$

for which the coefficients are alternating in sign and converging. If intrinsic defects form and migrate on the atomic vibration time-scale this series is expected to poorly converge. In this work we exclude any system configurations in which Frenkel defects have spontaneously formed, for example at the melting point, in order to ensure well-converged thermodynamic integrations of the anharmonic free energy of defect-free ZrC_x.
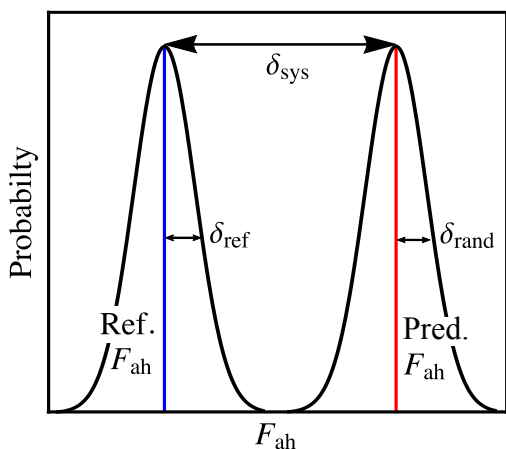

Figure 3. Schematic relating precision (δ_rand) and accuracy (δ_sys) errors in our predicted F_ah value to a reference F_ah value. The benchmark value is from TU-TILD[4] and is converged to $\delta_\text{ref} = 0.1$ meV/at. Precision δ_rand can easily be reduced to 0.1 meV/at or less, minimizing the systematic potential error. δ_sys is the main challenge.

Errors in predicting F_ah are considered from two primary sources, namely statistical convergence and a systematic potential error. The DFT benchmark also has a small convergence error which is accounted for, but other sources of error, such as from DFT exchange-correlation, electron-phonon scattering, and other quantum effects beyond the harmonic approximation, are beyond the scope of this paper. The three countable contributions are shown schematically in Fig. 3, and give the total expected error of

$$\delta = \sqrt{\delta_\text{sys}^2 + \delta_\text{rand}^2 + \delta_\text{ref}^2}. \tag{6}$$

Precision error δ_rand arises from evaluating an observable from a finite number of samples in the MEAM MD, and the systematic error δ_sys is due to the energy difference

between a MEAM potential and DFT. The convergence error in the benchmark F_ah value from TU-TILD is $\delta_\text{ref} = 0.1$ meV/at.[4]

The statistical convergence δ_rand is computed using stratified systematic sampling.[27] In the simulation the precision error scales as[28]

$$\delta_\text{rand} \sim \sqrt{\frac{\sigma_\lambda^2}{N_\lambda} \frac{\tau}{t}}, \tag{7}$$

where N_λ is the number of integration path points sampled (with $N_\lambda = 11$ in our case), σ_λ is the norm of ΔE standard deviations, τ is the ΔE autocorrelation time (ca. 11 fs, see Appendix), and t is the simulation time (ca. 0.1 ns). We emphasise that the TI method described differs from other approaches in that statistical convergence is not accuracy-limiting, for example, nanosecond simulations can comfortably be performed in a day on a low-performance computing platform.

The systematic potential error in the anharmonic free energy δ_sys can be computed by thermodynamic integration

$$\delta_\text{sys} = \int_0^1 d\lambda \, \langle E^\text{DFT}(\mathbf{R},\, V) - E^\text{MEAM}(\mathbf{R},\, V) \rangle_\lambda. \tag{8}$$

δ_sys is the primary error source in the method we describe to compute F_ah. Accuracy benchmarks in Sec. III A show δ_sys can be sufficiently controlled to satisfy 1 meV/at bulk convergence across the $F_\text{ah}(V,\, T)$ surface.

E. Technical details

Periodic plane-wave DFT calculations were performed using the VASP software,[29,30] with the local density approximation (LDA) exchange-correlation functional.[31] The projector-augmented wave (PAW) method was used,[32] with $4s$- and $4p$-Zr electrons included as valence states.

E_0 was computed on a mesh of 11 volumes, and at each volume the internal coordinates have been relaxed to give residual forces under 10^{-6} eV/Å. Self-consistent field (SCF) total energies and energy eigenvalues have been resolved to 10^{-9} eV. Methfessel-Paxton smearing has been used with a width of 0.1 eV.[33] The kinetic energy cutoff has been set to 700 eV and k-point mesh $12 \times 12 \times 12$ for the $2 \times 2 \times 2$ supercell. The E_0 vacancy formation energy contribution is extrapolated to the dilute limit, using data points from $\text{Zr}_{32}\text{C}_{31}$, $\text{Zr}_{108}\text{C}_{107}$ and $\text{Zr}_{256}\text{C}_{255}$.

For the quasiharmonic Helmholtz free energy F_qh, the kinetic energy cutoff has been set to 700 eV and k-point mesh to $6 \times 6 \times 6$ for the $2 \times 2 \times 2$ supercell. Phonons were calculated using the small displacement supercell method with the PHONOPY code.[34] At each of the 11 $2 \times 2 \times 2$ supercells that span the range of lattice parameters [4.575, 4.875] Å, sets of 18 displacements were made for $\text{Zr}_{32}\text{C}_{31}$ and sets of four displacements for $\text{Zr}_{32}\text{C}_{32}$. The phonon q-points were sampled by a mesh of $25 \times 25 \times 25$ points for the $2 \times 2 \times 2$ supercells.

The electronic Helmholtz free energy $F_{el}(V, T)$ has been calculated using the Mermin finite-temperature formulation of DFT,[35] on a mesh of 10 temperatures and 8 volumes sampled between $V_{eq}(T = 0\,\mathrm{K})$ and $V_{eq}(T_m)$. Electron states are self-consistent to at least 10^{-7} eV/atom. We used 384 bands, which was sufficient to span all states with partial occupation up to the melting point T_m. A kinetic cutoff energy of 700 eV was used, with k-point sampling at $12 \times 12 \times 12$ for the $2 \times 2 \times 2$ supercells.

The electron-vibration Helmholtz free energy $F_{el\text{-}vib}(V, T)$, has been calculated from low-converged MD configurations that are subsequently up-sampled, as in the proceedure recently performed for a number of transition metals.[36] The electronic free energy is calculated for each MD configuration, using the Mermin formulation at electronic temperature corresponding to the MD ensemble temperature. At each volume-temperature mesh point, the electronic free energies are averaged over the ensemble configurations, and referenced to the perfect crystal, in order to find the electron-vibration coupling contribution to the Helmholtz free energy.

The anharmonic Helmholtz free energy $F_{ah}(V, T)$ was determined using a mesh of six temperatures and five volumes. Temperatures span $0\,\mathrm{K}$ to T_m, and volumes $V_{eq}(T = 0\,\mathrm{K})$ to $1.15 V_{eq}(T_m)$. Potentials were fitted to MD configurations from DFT that used a 700 eV cutoff and k-point sampling mesh of $6 \times 6 \times 6$ for the $2 \times 2 \times 2$ supercell.
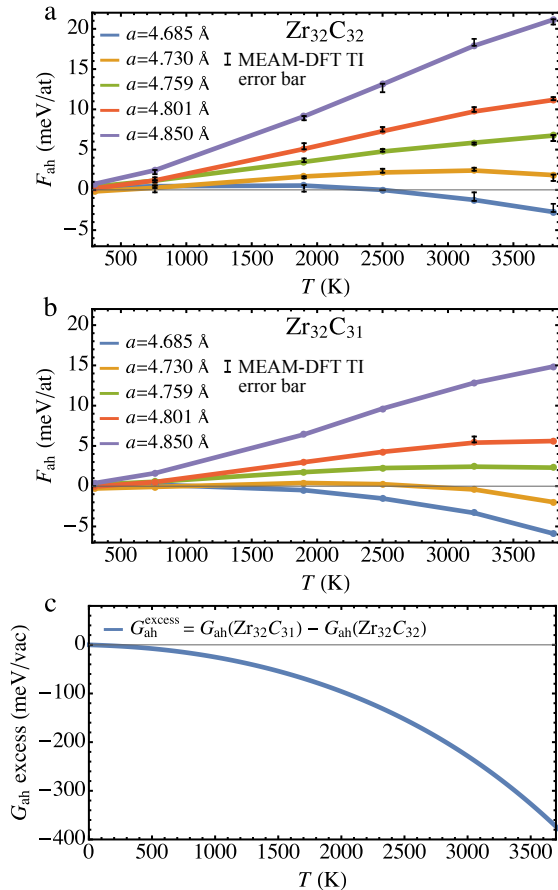
III. BENCHMARKS

A. Accuracy and precision

Figure 4. a) Anharmonic Helmholtz free energy F_{ah} for perfect $\mathrm{Zr}_{32}\mathrm{C}_{32}$. b) F_{ah} for $\mathrm{Zr}_{32}\mathrm{C}_{31}$. In each figure, error bars represent the deviation of F_{ah} from a DFT TI method reference (TU-TILD[4]). c) Vacancy anharmonicity, specified as an excess Gibbs free energy at ambient pressure.

Our MEAM TI approach predicts $F_{\mathrm{ah}}(V, T)$ to within a target accuracy of 1 meV/at compared to DFT endpoint TI. This is demonstrated in Fig. 4a. Perfect $\mathrm{Zr}_{32}\mathrm{C}_{32}$ is shown with error bars (TU-TILD reference) for 25 volumes and temperatures up to the melting point. $F_{\mathrm{ah}}(V, T)$ energies are converged to sufficient precision that the error bars are in effect systematic potential error bars. The mean absolute error (MAE) is 0.5 meV/at, with a mean signed deviation of 0.05 meV/at. The MAE values at the lattice parameters {4.685, 4.730, 4.759, 4.801, 4.850} Å are {0.72, 0.38, 0.35, 0.46, 0.64} meV/at. Errors resolved at the temperatures {760, 1900, 2500, 3200, 3805} K have the MAE values {0.39, 0.48, 0.43, 0.59, 0.66} meV/at.

On the basis of adequately small errors for bulk ZrC, we propose using the MEAM thermodynamic integration approach for more complex systems. In this regard $\mathrm{Zr}_{32}\mathrm{C}_{31}$ is a useful test case for two reasons. The carbon-vacancy introduces complexity in terms of physical interactions. It removes inversion symmetry at sites around the vacancy, so there are terms in the energy of odd-order in atomic displacements, previously excluded by symmetry in perfect ZrC. Secondly, making free energy predictions per vacancy increases computational complexity considerably due to the nature of statistical error scaling for TI predictions on a per-vacancy basis.

For the vacancy system $\mathrm{Zr}_{32}\mathrm{C}_{31}$, $F_{\mathrm{ah}}(V, T)$ is shown in Fig. 4b. Obtaining comparable DFT TI values for systems with vacancies like $\mathrm{Zr}_{32}\mathrm{C}_{31}$ is prohibitively expensive in general but we have computed a DFT benchmark for $\mathrm{Zr}_{32}\mathrm{C}_{31}$ at $a = 4.801$ Å and $T = 3200$ K. The MEAM thermodynamic integration is found to overestimate the DFT TI reference F_{ah} value by 0.4 meV/at, which is comparable to the MAE in the perfect bulk ZrC.

B. Computational cost

In Table I timings are reported for the MEAM-based TI in this work and TU-TILD (DFT) calculations. Both methods compute the $F_{\mathrm{ah}}(V, T)$ surface across 25 mesh points for a $\mathrm{Zr}_{32}\mathrm{C}_{32}$ test case. The MEAM approach does not have DFT TI and DFT TI up-sampling steps, which account for the majority of the TU-TILD F_{ah} cost. In the MEAM approach, the main CPU-time overhead is high-quality DFT calculations on selected MD configurations, which are used create data to fit the MEAM potentials. Furthermore the time required to optimize[25] the MEAM potential with a large fitting set is substantially longer (ca. ×10), compared to a MEAM potential used in the intermediate TI steps in TU-TILD. Despite this the former scheme still gains a factor of at least ×10 in efficiency overall due to having no DFT TI or DFT TI up-sample. For anharmonic predictions where 1 meV/atom convergence is sufficient, the TI method described in this work is likely to be a cost effective choice for metals and alloys. It would be of interest to compare the efficiency of less specialised machine learned potentials[7,37,38] to the MEAM type as applied here, in terms of parameter fitting time, required training DFT data and potential compute time.

IV. APPLICATION OF TI METHOD TO ZRC$_x$

A. The character of anharmonicity in ZrC_x

Prior to discussing the substoichiometric crystal, consider the anharmonic contribution to the Helmholtz free energy of perfect ZrC, shown in Fig. 4a. $F_{\mathrm{ah}}(V, T)$ in $\mathrm{Zr}_{32}\mathrm{C}_{32}$ tends to be positive and increase with temperature. This is because the anharmonic phase space has

Table I. Computer resources consumption for $\mathrm{Zr_{32}C_{32}}$ test case calculation of $F_{\mathrm{ah}}(V,T)$ on a 5×5 $\{V_i, T_i\}$ mesh. Timings listed in CPU core-hours and quasiharmonic-free-energy-job units ($1/t(F_{\mathrm{qh}})$) using the reference value $t(F_{\mathrm{qh}}) = 4800$ core-hours.

Contributions	$t(F_{\mathrm{ah}})$ (core-hrs)		$t(F_{\mathrm{ah}})/t(F_{\mathrm{qh}})$	
	this work	TU-TILD	this work	TU-TILD
Fit set DFT MD	10^4	10^3	10	0.6
MEAM fitting	10^3	10^2	0.4	0.03
MEAM TI	10^3	10^2	0.4	0.1
DFT TI	-	10^5	-	110
DFT up-sampling	-	10^4	-	6
Total	10^5	10^6	11	117

a smaller volume within a given potential energy surface, giving a positive anharmonic free energy term. A positive anharmonic contribution is similarly observed in other extended systems,[3,39,40] and is expected to be dependent on the presence of inversion symmetry.

$F_{\mathrm{ah}}(V,T)$ in $\mathrm{Zr_{32}C_{32}}$ increases with volume expansion. This contrasts with the result of harmonic force constants, which typically become softer under tensile strain, increasing the entropy and *decreasing* the free energy. In terms of effective frequencies in perfect ZrC, quasi-harmonicity reduces frequencies with volume expansion whereas anharmonicity in ZrC increases frequencies.

The anharmonic free energy of the $\mathrm{Zr_{32}C_{31}}$ crystal is given in Fig. 4b. $F_{\mathrm{ah}}(V,T)$ for $\mathrm{Zr_{32}C_{31}}$ naturally appears similar to $\mathrm{Zr_{32}C_{32}}$, since most atoms in $\mathrm{Zr_{32}C_{31}}$ are fully coordinated, but the anharmonic free energy is less positive, for example, $F_{\mathrm{ah}}(V,T)$ is lower by approximately 4 meV/at at $a = 4.759$ Å and $T = 3800$ K.

To directly identify vacant-site anharmonicity we compute

$$F_{\mathrm{ah}}^{\mathrm{excess}} = F_{\mathrm{ah}}\left(\mathrm{Zr_{32}C_{31}}\right) - F_{\mathrm{ah}}\left(\mathrm{Zr_{32}C_{32}}\right), \quad (9)$$

which isolates the vacancy anharmonic contribution by cancelling common contributions in $\mathrm{Zr_{32}C_{32}}$ and $\mathrm{Zr_{32}C_{31}}$. The anharmonicity of a single vacant site $F_{\mathrm{ah}}^{\mathrm{excess}}$ is stronger and qualitatively different in character to the anharmonicity per site in F_{ah} for $\mathrm{Zr_{32}C_{32}}$. $F_{\mathrm{ah}}^{\mathrm{excess}}$ for example typically exceeds F_{ah} by more than an order of magnitude (approximately $\times20$), and $F_{\mathrm{ah}}^{\mathrm{excess}}$ is negative whereas F_{ah} is almost always positive.

As we are typically interested in ambient pressure thermodynamics, we can consider the nature of the excess anharmonic Gibbs free energy $G_{\mathrm{ah}}^{\mathrm{excess}}$ rather than Helmholtz $F_{\mathrm{ah}}^{\mathrm{excess}}$. $G_{\mathrm{ah}}^{\mathrm{excess}}(T)$ in Fig. 4c illustrates the strength and sign of vacancy anharmonicity at ambient pressure. The large negative values of $G_{\mathrm{ah}}^{\mathrm{excess}}$ at high temperature can be simply rationalized. At high temperature the change in thermal excursions, when atoms are near a vacancy, is larger than predicted by harmonic springs, so the entropy is greater and free energy less. In terms of the change in the anharmonic potential, the magnitude and sign of $G_{\mathrm{ah}}^{\mathrm{excess}}$ are attributed to terms

that start from third-order in the potential Taylor expansion, rather than fourth-order as in the perfect crystal with inversion symmetry.

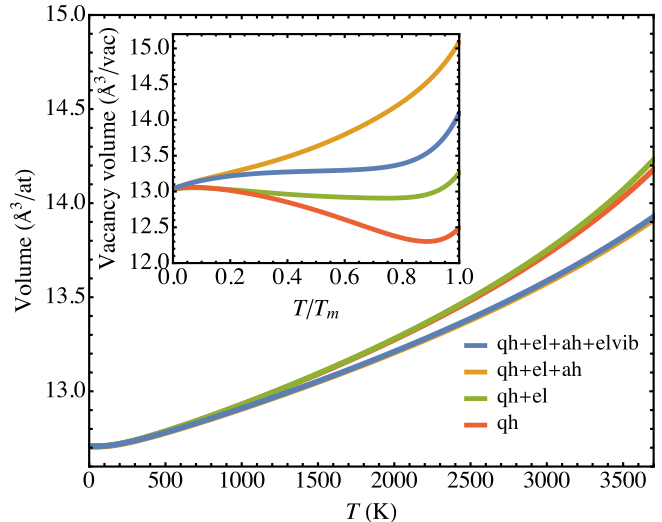## B. Vacancy volume, formation energy, and concentration

Figure 5. Thermal expansion of ZrC, $V(T)$, and inset, the vacancy formation volume $v^f(T)$.
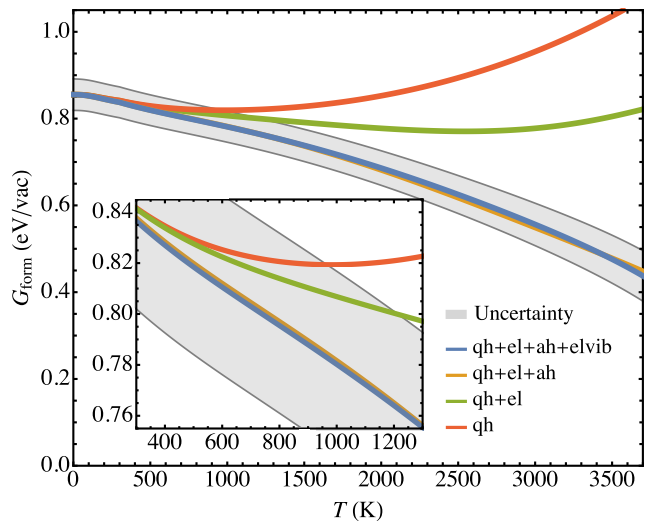

Figure 6. Gibbs free energy of carbon vacancy formation in ZrC versus temperature. *Inset*: The onset temperature of non-negligible anharmonicity. The error bar shown is for the TI method used to determine F_{ah}, from MEAM-DFT potential errors (assuming no cancellation between $\mathrm{Zr_{32}C_{32}}$ and $\mathrm{Zr_{32}C_{31}}$), and statistical convergence..

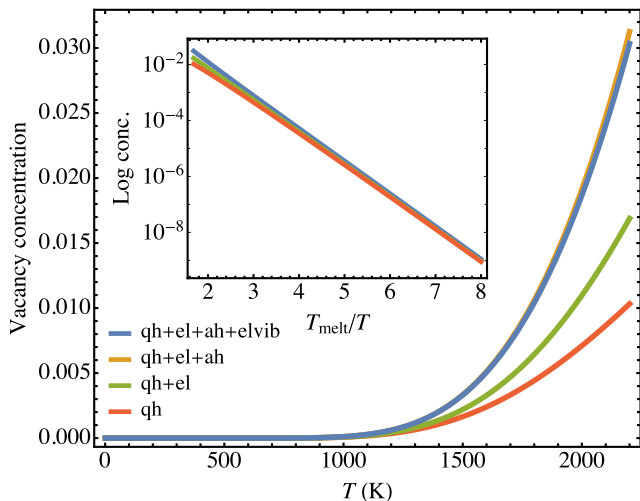The thermal expansion of the ZrC atomic volume (V) is shown in Fig. 5, alongside the vacancy formation vol-

Figure 7. Carbon vacancy concentration in ZrC_x versus temperature, up to a concentration of $c_{vac} = 1/32$. *Inset*: Log vacancy concentration versus T_m/T.

Table II. Concentration of vacancies (c_{vac}) in ZrC_x in carbon atom %.

T (K)	c_{vac} (C at. %)	
	this work	CALPHAD[41]
300	9×10^{-13}	1×10^{-15}
500	6×10^{-7}	1×10^{-9}
1000	0.01	0.003
1500	0.33	0.06
2000	1.87	0.33
2200	3.03	0.51

ume, $v^f(T) = \Omega - NV$, where Ω is the volume of an N atom defective ZrC_x cell. For ZrC, predictions at the qh+el+ah level of theory reproduce the thermal expansion reported in an earlier theoretical work,[4] while including electron-vibration coupling provides a small-to-negligible additional enhancement, evident in Fig. 5.

The volume of a vacant carbon site at $T = 298$ K is $v^f = 13.1$ Å^3/vac, which is +3.2 % or +0.40 Å^3/site larger than the corresponding atomic volume V for the perfect crystal. This means the lattice of a ZrC_x crystal initially expands for $x < 1$, with the lattice parameter a increasing by +0.001 Å from $x = 1$ to $x = 0.97$ in our 298-K calculations. This is at odds with recent measurements by Nakayama et al.[42] who report a monotonic, apparently linear trend, but is supported by other experimental work in which the ZrC_x lattice parameter is a concave function of carbon substoichiometry.[43–46] For example according to Sara,[43] the maximum volume occurs at composition $ZrC_{0.90}$ with an a value some +0.004 Å greater than in $ZrC_{0.98}$. To first order, the gradient is ca. +0.0011 Ang/C at. %, compared to +0.0005 Ang/C at. % in our work. It is important to stress that this is quite a subtle effect, and that it is temperature dependent. For $T > 2200$ K our computed lattice constant

decreases from $x = 1$ to $x = 0.97$.

As a final comment on thermal expansion, we note the temperature dependence of v^f is somewhat complicated (Fig. 5 inset). Quasiharmonic volume-dependent frequencies and electron-vibration coupling make the vacancy volume smaller generally, whereas anharmonic and electron thermal excitations increase it. In each instance, at high temperature such as $T \geq 0.75\,T_m$, these effects are comparable in size to the 0-K outward relaxation of the Zr atoms around the vacancy; ZrC bonds normal to the vacancy surface are squeezed by -0.08 Å compared to bonds of length $d(Zr-C) = 2.328$ Å in the perfect crystal.

The energy to form a carbon vacancy in ZrC is considered in terms of a Gibbs free energy computed as follows:

$$G_{form} = G(Zr_{32}C_{31}) + \mu(C) - G(Zr_{32}C_{32}). \quad (10)$$

In this expression $\mu(C)$ is the chemical potential to remove an atom of carbon from ZrC and place it in a carbon reservoir. The reference state of carbon is taken to be that of graphite, which for 0 K to 298 K, is computed by quasiharmonic DFT for diamond, with a 0-K experimental correction to graphite. At higher temperatures, the experimental parameterization of the graphite free energy is used, in the form of the Gustafson[47] assessment. This provides a diamond chemical potential that includes all contributions (e.g. anharmonicity) and is consistent with the DFT calculated free energies, while avoiding expensive calculations for graphite. Further details and an expression for $\mu(C)$ are given in the Appendix.

The error bar in $G_{form}(T_m)$, due to TI statistical precision and MEAM systematic potential error, is 60 meV/vacancy. This value assumes no cancellation in the systematic potential error between $Zr_{32}C_{32}$ and $Zr_{32}C_{31}$, and is therefore an upper limit. As most sites in $Zr_{32}C_{31}$ are fully-coordinated and bulk-like, and the same MEAM potential is used to describe perfect $Zr_{32}C_{32}$ and $Zr_{32}C_{31}$, partial cancellation of the systematic potential error is expected. In the limit of anharmonicity being a site-localised property, the systematic potential error would only arise from the six under-coordinated nearest-neighbours to the vacancy in $Zr_{32}C_{31}$, and the corresponding non-matching seven sites in $Zr_{32}C_{32}$. In this case the total error is less than 10 meV/vacancy.

$G_{form}(T)$ is shown in Fig. 6, including quasiharmonic, electronic, electron-vibration, and anharmonic contributions. Above ca. 1000 K ($T/T_m \approx 0.3$) the anharmonic contribution can no longer be regarded as negligible, and above ca. 2000 K ($T/T_m \approx 0.5$) accounting for anharmonicity is critical to qualitatively describe the ZrC vacancy formation energy. With respect to a quasiharmonic reference, Fig. 6 shows that electronic entropy lowers the formation energy, and that anharmonicity substantially lowers the Gibbs formation energy further, while the electron-vibration contribution is much smaller. In the final predictions, which include the quasiharmonic, electronic, electron-vibration, and anharmonic effects, G_{form} is almost linear in temperature, and decreases by approximately 10 meV with every increase in temperature by

1000 K. This rate of decrease is similar to reports in other materials such as aluminum and nickel.[3,48]

The vacancy concentration in contact with graphite is computed with an ideal solution model

$$\frac{c_{\text{vac}}}{1 - c_{\text{vac}}} = \exp\left(-\frac{G_{\text{form}}}{k_B T}\right), \quad (11)$$

and is shown in Fig. 7. Anharmonicity favors vacancy formation by making G_{form} smaller, increasing c_{vac} by a factor of two compared to predictions at the $F_{\text{qh}} + F_{\text{el}}$ level. The effect of electron-vibration coupling on vacancy concentration is marginal in this material. Specific values of $c_{\text{vac}}(T)$ are shown in Table II, up to a temperature of $T = 2200$ K, which is when the predicted concentration reaches our operative dilute limit of one vacancy per supercell ($c_{\text{vac}} = 1/32$ for ZrC_x).

In Table II the c_{vac} values from the CALPHAD assessment are consistently lower than our c_{vac} values.[41] Despite the power of the CALPHAD method for ZrC_x,[41] uncertainties can arise from insufficient experimental data, and the limitations that exist due to the non-physical interaction terms the methodology assumes. At $T = 2000$ K the CALPHAD value is $c_{\text{vac}} = 0.3$ C at. %,[41] compared to $c_{\text{vac}} = 1.9$ C at. % in this work.

Our predictions have quantum mechanical many-body errors from the LDA exchange-correlation treatment we use to describe ZrC. While the GGA has been shown to be less suitable to describe ZrC at high temperature than LDA,[4] it is instructive to consider the vacancy formation energy from both exchange-correlation treatments, in order to gauge sensitivity. At $T = 0$ K the GGA vacancy formation energy is less than the LDA value by some 0.2 eV/vacancy (without zero-point corrections and dilute limit supercell extrapolation), indicating a GGA predicted concentration is greater. Quantitative predictions of the non-local quantum many-body error at high temperature is beyond the scope of this work, but experience suggests that the LDA and GGA (PBE) functionals bracket the exact result.[2]

In this work we confine our predictions to temperatures at which concentrations do not exceed one vacancy per supercell. This should minimize lattice many-body errors, however we note that vacancy-vacancy interactions are expected to be mainly repulsive,[49] and that other entities on the carbon sub-lattice such as Frenkel defects will decrease the vacancy configuration space.[50] These effects are expected to moderate c_{vac}, to values lower than ideal, to an extent that increases with temperature.

C. Free energy and heat capacity of ZrC_x

The basic excitation mechanisms that determine the thermal properties of ZrC_x are discussed relative to a quasiharmonic reference system. Formulae for the referenced Gibbs free energies at ambient pressure, G_{el}, $G_{\text{el-vib}}$, G_{ah} and G_{vac}, are listed in the Appendix. Each is shown as a function of temperature in Fig. 8.
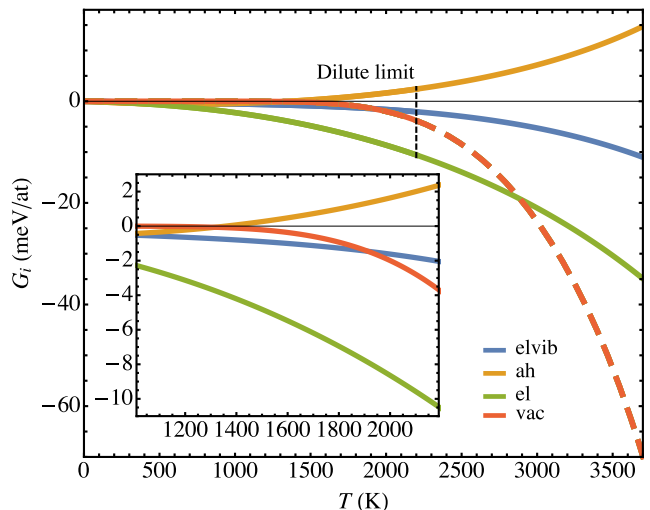

Figure 8. ZrC_x Gibbs free energy contributions with respect to a quasiharmonic reference. *Inset*: Gibbs free energy shown from 1200 K to the dilute vacancy concentration (1/32) at 2200 K.
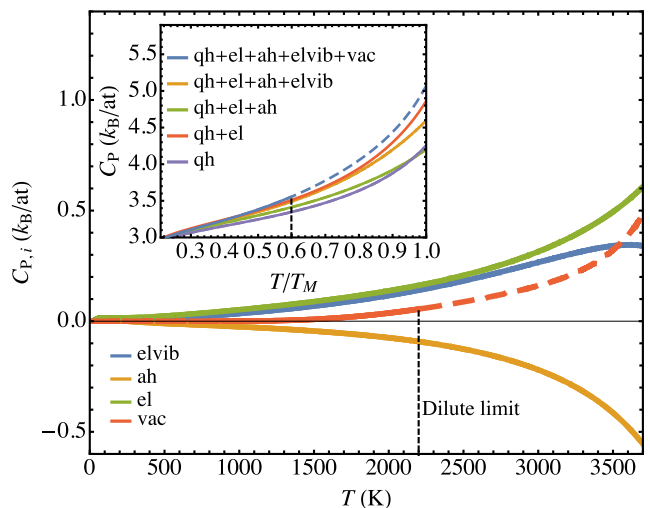

Figure 9. ZrC_x constant-pressure heat capacity relative to a quasiharmonic reference. *Inset*: heat capacity at different levels of theory, shown from the Debye temperature ($C_p(T_{\text{Debye}}) = 3\ k_B$) to the melting point ($T_m = 3700$ K).

At high temperature the magnitude of the electron-vibration contribution to the Gibbs free energy is less than the anharmonic contribution, which is in turn less than the electronic contribution. Partial cancellation occurs as G_{el} and $G_{\text{el-vib}}$ are negative whereas G_{ah} is positive in this material. The vacancy contribution G_{vac} is the smallest of the four contributions up to 1900 K, but beyond the dilute vacancy concentration temperature of $T = 2200$ K, G_{vac} increases considerably. Extrapolating to higher temperatures, the vacancy contribution appears to become the largest of all above 3000 K. Note however that above $T = 2200$ K the value of G_{vac} is presented

as indicative only, and is represented in Fig. 8 with a dashed line, as it exceeds the thermodynamic limitations of our dilute solution model.

The different ZrC_x heat capacity contributions, relative to the quasiharmonic system, are shown in Fig. 9. The anharmonic term $C_{\mathrm{P,ah}}$ is negative and the electronic one $C_{\mathrm{P,el}}$ is positive, with each similar in absolute value near T_m. It is somewhat interesting to consider the extent to which $C_{\mathrm{P,ah}}$ and $C_{\mathrm{P,el}}$ cancellation is coincidental in ZrC_x or a manifestation of a generic feature. In ordinary metals $C_{\mathrm{P,el}}$ is *a priori* positive, and for ordinary high-symmetry crystals, $C_{\mathrm{P,ah}}$ is negative at high temperature.[39] Some cancellation of $C_{\mathrm{P,ah}}$ and $C_{\mathrm{P,el}}$ is therefore regarded to be likely in conducting systems such as the refractory ceramic ZrC. However, as the extent of cancellation depends on the magnitude of each contribution, for which we are unaware of a direct physical relation, we conclude cancellation is mostly coincidental.

In ZrC_x, $C_{\mathrm{P,ah}}$, $C_{\mathrm{P,el}}$, and $C_{\mathrm{P,el\text{-}vib}}$ are all individually larger than the vacancy contribution. For example at 2200 K, $C_{\mathrm{P,vac}} = +0.06$ $k_{\mathrm{B}}/\mathrm{atom}$, compared to $C_{\mathrm{P,el}} = +0.16$ $k_{\mathrm{B}}/\mathrm{atom}$, $C_{\mathrm{P,el\text{-}vib}} = +0.14$ $k_{\mathrm{B}}/\mathrm{atom}$ and $C_{\mathrm{P,ah}} = -0.09$ $k_{\mathrm{B}}/\mathrm{atom}$. Extending the vacancy model beyond the dilute limit with the dashed line in Fig. 9 indicates $C_{\mathrm{P,vac}}$ is comparable to the positive contributions of $C_{\mathrm{P,el}}$, and $C_{\mathrm{P,el\text{-}vib}}$ near T_m. Consequently we suggest the physical origin of the steep increase in heat capacity in ZrC_x[51] is a combination of electron thermal excitations, electron-vibration coupling, and structural excitations on the carbon sub-lattice, predominantly the constitutional carbon vacancies that have been the focus of this paper, although there are also stoichiometry conserving intrinsic carbon Frenkel defects, which are beyond the scope of this work but discussed elsewhere.[50,51] Finally, it is interesting to note that while anharmonicity is the only term that suppresses C_{P} in Fig. 9, the enhancing effects of the vacancy contribution are mainly due to the indirect effect of anharmonicity, insomuch as anharmonicity by lowering G_{form} enhances the population of vacancies (e.g. by a factor of $\times 2$ at 2200 K as shown in Fig. 7).

V. CONCLUSIONS

Thermodynamic integration approach

We have described a thermodynamic integration method to calculate the anharmonic free energy of a crystal to DFT accuracy. In our benchmark system of ZrC the approach achieves average target precision better than 1 meV/atom and 60 meV/vacancy, comparable to more expensive DFT-based thermodynamic integration schemes. The method transfers the burden of computation from converging random statistical errors to minimizing systematic potential errors. With sufficient care to minimize potential errors, precise calculations can be made to compute quantities such as the anharmonic vacancy formation energy, with modest computing resources.

Application to ZrC_x

Anharmonicity increases the concentration of vacancies in ZrC_x. At 2000 K vacancies increase from 1.1 to 1.9 C at. % due to explicit anharmonicity. The ZrC_x heat capacity is enhanced by electron thermal excitations, electron-vibration coupling and vacancies on the carbon sub-lattice, and suppressed by anharmonicity. For example, $C_{\mathrm{P,el}}(T = 2200\,\mathrm{K}) = +0.16$ $k_{\mathrm{B}}/\mathrm{atom}$, $C_{\mathrm{P,el\text{-}vib}}(T = 2200\,\mathrm{K}) = +0.14$ $k_{\mathrm{B}}/\mathrm{atom}$, $C_{\mathrm{P,vac}}(T = 2200\,\mathrm{K}) = +0.06$ $k_{\mathrm{B}}/\mathrm{atom}$, and $C_{\mathrm{P,ah}}(T = 2200\mathrm{K}) = -0.09$ $k_{\mathrm{B}}/\mathrm{atom}$. The sharp increase in the heat capacity at high temperature is attributed to electronic and electron-vibration effects along with the thermal excitation of structural defects.

VI. ACKNOWLEDGEMENTS

VII. REFERENCES

[*] t.mellan@imperial.ac.uk

[1] B. Grabowski, L. Ismer, T. Hickel, and J. Neugebauer, Phys. Rev. B **79**, 134106 (2009).

[2] B. Grabowski, S. Wippermann, A. Glensk, T. Hickel, and J. Neugebauer, Phys. Rev. B **91**, 201103(R) (2015).

[3] A. Glensk, B. Grabowski, T. Hickel, and J. Neugebauer, Phys. Rev. X **4**, 011018 (2014).

[4] A. I. Duff, T. Davey, D. Korbmacher, A. Glensk, B. Grabowski, J. Neugebauer, and M. W. Finnis, Phys. Rev. B **91**, 214311 (2015).

[5] A. I. Duff, T. A. Mellan, and M. W. Finnis, In preparation (2018).

[6] Y. Ikeda, B. Grabowski, and F. Körmann, Mat. Char. **147**, 464 (2018).

[7] B. Grabowski, F. Körmann, C. Freysoldt, A. I. Duff, A. Shapeev, and J. Neugebauer, submitted arXiv:1902.11230.

[8] Z. Wu and R. M. Wentzcovitch, Phys. Rev. B **79**, 104304 (2009).

[9] S. G. Moustafa, A. J. Schultz, and D. A. Kofke, J. Chem. Theo. Comp. **13**, 825 (2017).

[10] A. Purohit, A. J. Schultz, S. G. Moustafa, J. R. Errington, and D. A. Kofke, Mol. Phys. **116**, 3027 (2018).

[11] X. Zhang, B. Grabowski, T. Hickel, and J. Neugebauer, Comp. Mat. Sci. **148**, 249 (2018).

[12] D. Alfè, G. D. Price, and M. J. Gillan, Phys. Rev. B **64**, 045123 (2001).

[13] D. Alfè, M. J. Gillan, and G. D. Price, J. Chem. Phys. **116**, 6170 (2002).

[14] D. Alfè, G. D. Price, and M. J. Gillan, Phys. Rev. B **65**, 165118 (2002).

[15] G. Ackland, J Phys: Cond. Mat. **14**, 2975 (2002).

[16] I. Errea, B. Rousseau, and A. Bergara, Phys. Rev. Lett. **106**, 165501 (2011).

[17] O. Hellman, P. Steneteg, I. A. Abrikosov, and S. I. Simak, Phys. Rev. B **87**, 104111 (2013).

[18] B. Monserrat, N. D. Drummond, and R. J. Needs, Phys. Rev. B **87**, 144302 (2013).

[19] M. L. Klein and G. K. Horton, J. Low Temp. Phys. **9**, 151 (1972).

[20] R. W. Zwanzig, J. Chem. Phys. **22**, 1420 (1954).

[21] D. Frenkel and A. J. C. Ladd, J. Chem. Phys. **81**, 3188 (1984).

[22] S. G. Moustafa, A. J. Schultz, E. Zurek, and D. A. Kofke, Phys. Rev. B **96**, 014117 (2017).

[23] J. G. Kirkwood, J. Chem. Phys. **3**, 300 (1935).

[24] A. I. Duff, M. W. Finnis, P. Maugis, B. J. Thijsse, and M. H. F. Sluiter, Comp. Phys. Comm. **196**, 439 (2015).

[25] A. I. Duff, Comp. Phys. Comm. **203**, 354 (2016).

[26] "MEAMfit2 is an interatomic potential optimization package which can be obtained from STFC's Daresbury laboratory via the website https://www.scd.stfc.ac.uk/Pages/MEAMfit-v2.aspx,".

[27] J. M. Haile, *Molecular Dynamics Simulation: Elementary Methods* (Wiley, 1997) p. 79f.

[28] W. Janke, in *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*, Vol. 10 (John von Neumann Institute for Computing, Jülich, 2002) p. 423.

[29] G. Kresse and J. Furthmüller, Comp. Mat. Sci. **6**, 15 (1996).

[30] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).

[31] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[32] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).

[33] M. Methfessel and A. T. Paxton, Phys. Rev. B **40**, 3616 (1989).

[34] A. Togo and I. Tanaka, Scrip. Mat. **108**, 1 (2015).

[35] N. D. Mermin, Phys. Rev. **137**, 1441 (1965).

[36] X. Zhang, B. Grabowski, F. Körmann, C. Freysoldt, and J. Neugebauer, Phys. Rev. B **95**, 165126 (2017).

[37] B. Cheng, E. Engel, J. Behler, C. Dellago, and M. Ceriotti, PNAS **116**, 1110 (2019).

[38] J. Behler, Ang. Chem. **56**, 12828 (2017).

[39] A. Glensk, B. Grabowski, T. Hickel, and J. Neugebauer, Phys. Rev. Lett. **114**, 195901 (2015).

[40] X. Zhang, B. Grabowski, F. Körmann, A. Ruban, Y. Gong, R. Reed, T. Hickel, and J. Neugebauer, Phys. Rev. B **98**, 224106 (2018).

[41] A. Fernández Guillermet, J. Alloy Compd. **217**, 69 (1995).

[42] H. Nakayama, K. Ozaki, T. Nabeta, and Y. Nakajima, Mat. Trans. **58**, 852 (2017).

[43] R. Sara, J. Amer. Cer. Soc. **48**, 243 (1965).

[44] K. Schönfeld, H. Martin, and A. Michaelis, J. Adv. Cer. **6**, 165 (2017).

[45] Y. Katoh, G. Vasudevamurthy, T. Nozawa, and L. Snead, J. Nuc. Mat. **441**, 718 (2013).

[46] H. F. Jackson, J. D. D. Jayaseelan, D. Manara, P. Casoni, and W. E. Lee, J. Am. Ceram. Soc. **94**, 3561 (2011).

[47] P. Gustafson, Carbon **24**, 169 (1986).

[48] Y. Gong, B. Grabowski, A. Glensk, F. Körmann, J. Neugebauer, and R. C. Reed, Phys. Rev. B **97**, 214106 (2018).

[49] Y. Zhang, B. Liu, and J. Wang, Sci. Rep. **5**, 18098 (2015).

[50] T. A. Mellan, A. I. Duff, and M. W. Finnis, Phys. Rev. B **98**, 174116 (2018).

[51] A. I. Savvatimskiy, S. V. Onufriev, and S. A. Muboyadzhyan, J. Mater. Res. **32**, 1287 (2017).

APPENDIX

GIBBS FREE ENERGY CONTRIBUTIONS

To quantify different thermal excitations beyond the quasiharmonic level of theory, quasiharmonic-referenced Gibbs free energies are reported as

$$G_{ah} = \min_{V} \left[F_{ah} + F_{qh} + E_0 + pV \right] - \min_{V} \left[F_{qh} + E_0 + pV \right],$$

$$G_{el\text{-}vib} = \min_{V} \left[F_{el\text{-}vib} + F_{qh} + E_0 + pV \right] - \min_{V} \left[F_{qh} + E_0 + pV \right],$$

and

$$G_{el} = \min_{V} \left[F_{el} + F_{qh} + E_0 + pV \right] - \min_{V} \left[F_{qh} + E_0 + pV \right].$$

Similarly G_{vac} is the Gibbs free energy associated with a concentration of vacancies (c_{vac}) in ZrC_x, again referenced to the quasiharmonic system. This is defined by writing the total Gibbs free energy of ZrC_x at the full level of theory as $G = G^{\text{perf}} - c_{\text{vac}} k_{\text{B}} T$. In this expression $G^{\text{perf}} = \min_V [F_{\text{ah}} + F_{\text{el-vib}} + F_{\text{qh}} + F_{\text{el}} + E_0 + pV]$, and c_{vac} is the equilibrium concentration of vacancies, which has been computed from the Arrhenius ideal solution model introduced in Eqn. (11). To compute G_{vac}, the vacancy part ($G - G^{\text{perf}}$) is referenced to the quasiharmonic system:

$$G_{\text{vac}} = \left(G - G^{\text{perf}}\right) - \left(G_{\text{qh}} - G_{\text{qh}}^{\text{perf}}\right)$$
$$= -\left(c_{\text{vac}} - c_{\text{vac}}^{\text{qh}}\right) k_{\text{B}} T,$$

which is equivalent to the difference in equilibrium vacancy concentrations at the full and quasiharmonic levels of theory. Note for completeness, the quasiharmonic reference system terms are defined as follows: $G_{\text{qh}} = G_{\text{qh}}^{\text{perf}} - c_{\text{vac}}^{\text{qh}} k_{\text{B}} T$, with $G_{\text{qh}}^{\text{perf}} = \min_V [F_{\text{qh}} + E_0 + pV]$, and $c_{\text{vac}}^{\text{qh}}$ is the ideal solution model equilibrium concentration, with exponent G_{form} (Eqn. (10)) calculated at the quasiharmonic level.

GRAPHITE CHEMICAL POTENTIAL

The ZrC vacancy formation energy has been calculated with respect to a graphite chemical potential of the form

$$\mu(\text{C}) = \begin{cases} G_{\text{diamond}}(T) + (H_{\text{graphite}} - H_{\text{diamond}}) & 0\,\text{K} < T \leq T_{\text{stn}} \\ \sum_{-3 \leq i \leq 2} a_i(T^i - T_{\text{stn}}^i) + a_3 \left[T \ln(T) - T_{\text{stn}} \ln(T_{\text{stn}})\right] + G_{\text{diamond}}(T_{\text{stn}}) + (H_{\text{graphite}} - H_{\text{diamond}}) & T_{\text{stn}} < T \leq T_m \end{cases}.$$

The coefficients for $T > T_{\text{stn}}$, which is the CALPHAD standard state temperature $T_{\text{stn}} = 298.15$ K, are set according to the Gustafson experimental free energy parameterization:[47]

$$a_{[-3, 3]} = \left\{1.2 \times 10^{10}, -2.643 \times 10^8, 2562600, -17369, 170.73, -4.723 \times 10^{-4}, -24.3\right\}.$$

At low temperatures ($T \leq 298.15$ K) where the graphite parameterization is unavailable, $\mu(\text{C})$ is continued using a DFT-calculated diamond potential, $G_{\text{diamond}}(T)$. This quasiharmonic diamond potential at low-temperature is transformed to a graphite chemical potential by a correction equal to the 0 K enthalpy difference $(H_{\text{graphite}} - H_{\text{diamond}}) = -0.03$ eV/atom.

with time-step $\Delta t = 1$ fs, and autocorrelation function, $A(k)$, given by

$$A(k) = \frac{\langle \Delta U_i \Delta U_{i+k} \rangle - \langle \Delta U_i \rangle \langle \Delta U_i \rangle}{\langle \Delta U_i^2 \rangle - \langle \Delta U_i \rangle \langle \Delta U_i \rangle}.$$

The calculated correlation time ranges from $9 - 13$ fs, depending weakly on the arguments of $\tau(T, V, \lambda)$.

TI QUADRATURE ERROR

The quadrature error in F_{ah} as a function of the number of integral sampling points (λ_i values) is shown in Fig. 10. At low temperatures when the system is nearly harmonic, the integrand $\langle \partial_\lambda E_{\text{mix}}(\lambda) \rangle_\lambda$ is small and almost independent of λ, and therefore a large number of λ_i samples is unnecessary. At high temperature when $\langle \partial_\lambda E_{\text{mix}}(\lambda) \rangle_\lambda$ is curvier (see non-linearity in Fig. 2), sufficient sampling of the integrand is critical to obtain sub-meV/at numerical precision. F_{ah} is determined in this work by sampling $\langle \partial_\lambda E_{\text{mix}}(\lambda) \rangle_\lambda$ at 10 intervals or 11 points ($\lambda_i = i/10$). The associated error shown in Fig. 10 is less than 0.1 meV/at for $T \leq 3200$ K, and ca. 0.2 meV/at at T_m.
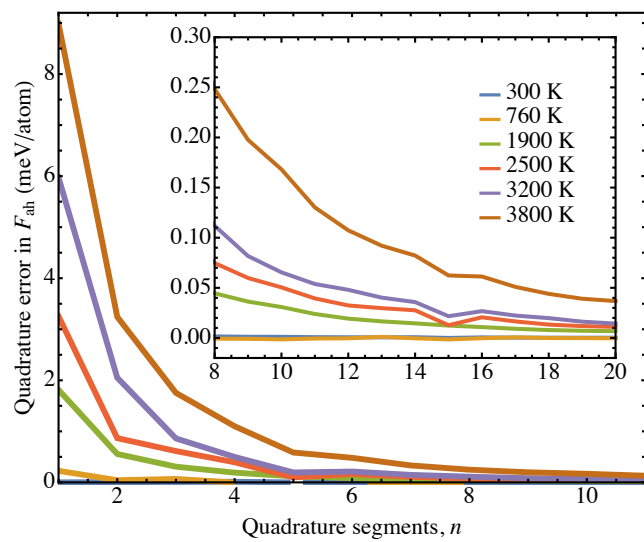
CORRELATION TIME

The correlation time τ, which is used to estimate error scaling and determine statistical precision using stratified systematic sampling, is the integrated correlation time[28]

$$\tau(T, V, \lambda) \equiv \tau_{\Delta U, \text{int}}.$$

$\tau_{\Delta U, \text{int}}$ is estimated by[28]

$$\tau_{\Delta U, \text{int}} = \left(\frac{1}{2} + \sum_{k=1}^{N} A(k)\right) \Delta t,$$

Figure 10. Thermodynamic integration error in F_{ah} as function of the number of quadrature segments n in $\lambda_i = i/n$.