# SCIENTIFIC REPORTS

**OPEN**

# No system-size anomalies in entropy of bcc iron at Earth's inner-core conditions

Andrew J. Schultz, Sabry G. Moustafa & David A. Kofke

New molecular modeling data show that the entropy of bcc iron exhibits no system-size anomalies, implying that it should be feasible to compute accurate free energies of this system using first-principles methods without requiring a prohibitively large number of atoms. Conclusions are based on rigorous calculations of size-dependent free energies for a Sutton-Chen model of iron previously fit to ab initio calculations, and refute statements recently appearing in the literature indicating that the size of the simulation cell is critical for stabilization of the bcc phase.

Recently, data from molecular simulations of iron at Earth's inner-core conditions have been reported in ref.[1] (hereafter referred to as (I)). A key focus of (I) is the stability of the bcc phase relative to hcp at these conditions. A major claim made there is that a diffusion mechanism stabilizes bcc, and that simulations of very large systems are needed for this effect to be realized. Several observations are made in (I) to provide evidence for this conclusion. Primarily it is found that isobaric molecular dynamics (MD) simulations of atomistic models of iron using density functional theory (DFT) spontaneously transition from bcc to hcp when simulated using a small number of atoms ($N = 432$), but for larger systems ($N = 1024$) such transitions are not observed for the duration of 18-ps simulations at 7000 K. However, at lower temperature (6000 K), the 1024-atom simulations also spontaneously transition to hcp, and it is argued that it is not possible to determine whether this is still a system-size effect, or the true behavior for this temperature. *Ab initio* simulations of larger systems needed to address this are not presently feasible. Examination of configurations and trajectories in (I) shows concerted diffusive processes, and the authors of (I) assert that this behavior contributes to the entropy and thereby stabilizes bcc; moreover they state that such processes cannot be supported by small systems without destroying their structure. This then, according to (I), is the origin of the strong size dependence of the bcc stability.

To examine this phenomenon further, the authors of (I) parameterized a Sutton-Chen (SC) embedded-atom model by fitting to DFT energies for select configurations of bcc and hcp phases observed in 1024-atom simulations at 6000 K. Energies and forces from SC are much less computationally expensive to evaluate compared to DFT, so the SC system could be simulated for much larger system sizes; results from simulations of up to $16 \times 10^6$ atoms were reported in (I). The same behavior is observed in the SC calculations as seen for DFT: the system spontaneously transitions from bcc to hcp when simulating smaller systems and lower temperatures, while this transition is not observed over the duration of simulations of systems with greater numbers of atoms. The bcc phase is observed to persist for at least 200 ps in 65,536-atom simulations for temperatures as low as $T = 5500$ K (but not 5000 K). We have independently performed a more limited set of simulations of the SC model and have observed the same general behaviors as described in (I).

To put the observations on a thermodynamic footing, the authors of (I) collected averages of the SC hcp and bcc dynamics and used these data in an approximate method, based on analysis of the velocity autocorrelation function (VACF), to estimate the entropy of each phase. Their results for the bcc entropy per mole $S_{bcc}/N$ show a remarkable dependence on system size (Fig. 1), which the authors of (I) use to explain the increase in mechanical stability of bcc with $N$. The strong variation of the entropy with system size (even up to $N = 20,000$) violates the expected linear scaling that is characteristic of an extensive thermodynamic quantity (in contrast, the hcp entropy per mole varies little with $N$). This unusual behavior is ascribed to the enhanced bcc diffusion observed in both the DFT- and the SC-based studies, which is manifested only for very large $N$. The implication is that simulations of small systems are unable to capture the essential phenomena that govern the thermodynamic stability of bcc iron.

Department of Chemical and Biological Engineering, University at Buffalo, The State University of New York, Buffalo, New York, 14260-4200, USA. Correspondence and requests for materials should be addressed to D.A.K. (email: kofke@buffalo.edu)
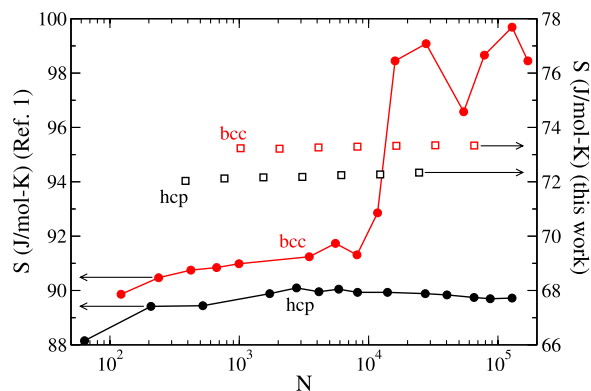
**Figure 1.** Entropy for bcc (red) and hcp (black) phases of iron as a function of simulated-system size at 360 GPa and 7000 K. Filled symbols joined by lines (left axis) are results from (I); open symbols (right axis) are results computed here, with estimated stochastic + systematic errors that are smaller than symbol size. Note that left and right axes are shifted with respect to each other, but cover the same range (12 J/mol-K).

The issue of a non-extensive entropy for moderate $N$ is important, because a need for very large $N$ to capture bcc behavior correctly would preclude evaluation of the entropy using an accurate DFT model, which is too computationally expensive for the large system sizes prescribed by (I). If true, this means that we may give up any hope of computing an accurate phase diagram for iron at extreme conditions from first-principles considerations.

Of course, diffusion *per se* does not contribute to the free energy, which is an equilibrium thermodynamic property that can be defined and computed without any reference to dynamic behavior. Hence, it should be possible to demonstrate the anomalous entropy scaling just as well through rigorous free-energy calculations using any convenient simulation method. For force-field models (such as SC) it is possible to evaluate the free energy very accurately (within any approximation inherent in the molecular model), enabling the observations made in (I) to be rigorously tested. It is still a challenging calculation however, because at the pressures of interest bcc is mechanically (and thus thermodynamically) unstable at $T = 0$ K, so we cannot apply a conventional approach based on integration from low temperature[2]; further, results from methods using integration from a noninteracting cell model or Einstein lattice[3] should work but might be viewed with some doubt because they completely suppress diffusion along part of the integration path, and the switch from non-diffusive to diffusive behavior could be problematic.

Nevertheless, we have completed such calculations, generating very accurate entropies for the SC model developed in (I), for different system sizes $N$. The approach we take to calculate the $N$-dependent Helmholtz free energy is to perform successive calculations in which we compute the difference $\Delta A(N)$ in free energy upon doubling of the system size, $\Delta A(N) \equiv A(T, 2V, 2N) - A(T, V, N)$. This type of calculation has been proposed previously[4–6] as a means to obtain the absolute free energy, by coupling the result with the assumption that the free energy is a first-order homogeneous function of $V$ and $N$: $A(T, 2V, 2N) = 2A(T, V, N)$ (i.e., the free energy is an extensive thermodynamic quantity). In the present case, our primary interest is to test whether the free energy is extensive for small-to-moderate values of $N$, so instead we perform this calculation for successively larger systems, repeatedly doubling the system size, and we examine whether the successive differences are consistent with an extensive free energy. The treatment is appealing because the calculated free energy differences coincide with the independent variable of interest ($N$), and moreover the system is at all stages stable (or at least metastable) and free to undergo some diffusive processes.

We find that the free energy is indeed extensive, and there is no anomalous behavior of $S_{bcc}$ with system size. Results are presented in the next section, followed by a discussion, and a description of the methods used.

## Results

Our calculations yield absolute free-energy values including all contributions (i.e. lattice, harmonic, and anharmonic) due to motion of the nuclei according to the Sutton-Chen semi-empirical potential given in (I). Inasmuch as the SC potential was fit to DFT calculations using Fermi-Dirac occupation (smearing), results obtained from the model in principle include electronic excitation contributions, hence they do not need to be accounted for separately.

We performed system-doubling simulations to compute free energy differences $\Delta A(N_j)$ for $N_j \rightarrow 2N_j$, with $N_j$ as small as 192 atoms and as large as 32,768 (doubling to 65,536), as detailed in the methods section. The $\Delta A(N_j)$ values for bcc and hcp phases were computed from equation (16). Now, if $\Delta A(N_j)/N_j$ is independent of $N_j$, then growth of $A$ with $N$ (as given by successive increments in $\Delta A$) is linear in $N$, and hence $A$ is extensive. Therefore, to gauge whether $A$ is extensive, we may examine $\Delta A(N_j)/N_j$ to see whether it is constant with $j$. Apart from this, we note also that for an extensive free energy, $A(N_j)/N_j = \Delta A(N_j)/N_j$, so as we examine the variation in $\Delta A(N_{j-1})/N_{j-1}$, it is useful to regard it in terms of $A(N_j)/N_j$ itself, thereby allowing comparison to the free energy values reported in (I). Thus, in Table 1 we report the values of $\Delta A(N_{j-1})/N_{j-1}$ obtained as described above, but we label them in the table as $A(N_j)/N_j$.

We compute the entropy $S$ from the Helmholtz free energy via

| | bcc | | | | hcp | | | |
|---|---|---|---|---|---|---|---|---|
| $N$ | $A/Nk_BT$ | $\langle U\rangle/Nk_BT$ | $S/Nk_B$ | $G/Nk_BT$ | $A/Nk_BT$ | $\langle U\rangle/Nk_BT$ | $S/Nk_B$ | $G/Nk_BT$ |
| $N_1$ | −17.4380(12) | −8.636(6) | 8.802(6) | 7.3330(12) | −17.3387(18) | −8.673(8) | 8.665(8) | 7.3051(18) |
| $N_2$ | −17.4323(13) | −8.623(3) | 8.809(3) | 7.3386(13) | −17.3435(18) | −8.688(6) | 8.656(6) | 7.3003(18) |
| $N_3$ | −17.445(2) | −8.6214(18) | 8.823(3) | 7.326(2) | −17.344(2) | −8.684(3) | 8.660(4) | 7.299(2) |
| $N_4$ | −17.448(5) | −8.616(2) | 8.831(5) | 7.323(5) | −17.349(2) | −8.685(3) | 8.664(4) | 7.295(2) |
| $N_5$ | −17.450(6) | −8.6223(12) | 8.827(7) | 7.321(6) | −17.348(3) | −8.6775(18) | 8.670(3) | 7.296(3) |
| $N_6$ | −17.443(9) | −8.6152(10) | 8.827(9) | 7.328(9) | −17.347(4) | −8.6838(16) | 8.663(4) | 7.297(4) |
| $N_7$ | −17.453(11) | −8.6189(8) | 8.834(11) | 7.318(11) | −17.346(4) | −8.6793(12) | 8.666(4) | 7.298(4) |

**Table 1.** Thermodynamic properties computed in this work. $N_j$ is the system size as defined in Table 3; $A$ is the Helmholtz free energy; $\langle U\rangle$ is the internal potential energy averaged in the $w=0$ limit; $S$ is the entropy; $G$ is the Gibbs free energy, assuming a pressure of 360 GPa. Numerals in parentheses indicate the estimated error (stochastic + systematic; see text) in the rightmost digit(s) of the reported value. All data are for $T = 7000$ K and specific volumes of 6.65 Å$^3$/atom for the bcc phase, and 6.616 Å$^3$/atom for hcp.

| Source | $\delta A/Nk_BT$ | $\delta\langle U\rangle/Nk_BT$ | $P\delta V/Nk_BT$ | $\delta H/Nk_BT$ | $\delta S/Nk_B$ | $\delta G/Nk_BT$ |
|---|---|---|---|---|---|---|
| This work | −0.107(15) | 0.0604(14) | 0.1272 | 0.1876(14) | 0.167(15) | 0.020(15) |
| Ref.[1] | | | | | (0.345, 0.962) | |

**Table 2.** Differences $\delta \equiv$ (bcc) − (hcp) in thermodynamic properties for the Sutton-Chen model defined in (I). "This work" indicates results as computed in this work, using property values for the systems of size $N_6$ for each phase. Entropy is reported in ref.[1] as a range of values, represented here by the pair of numbers in parentheses. $A$ is the Helmholtz free energy; $\langle U\rangle$ is the energy; $H$ is the enthalpy; $S$ is the entropy; $G$ is the Gibbs free energy, $P = 360$ GPa is the pressure, $V/N$ is the volume per atom (6.65 Å$^3$/atom for the bcc phase, and 6.616 Å$^3$/atom for hcp). Numerals in parentheses indicate the estimated error (68% confidence) in the rightmost digit(s) of the reported value.

| | bcc | | | | hcp | | | |
|---|---|---|---|---|---|---|---|---|
| $j$ | $N_j \rightarrow N_{j+1}$ | Supercell | $\lambda_m$ | $n_\lambda$ | $N_j \rightarrow N_{j+1}$ | Supercell | $\lambda_m$ | $n_\lambda$ |
| 0 | 512 → 1024 | 8 × 8 × (4 → 8) | 19.0 | 227 | 192 → 384 | 6 × 4 × (2 → 4) | 19.0 | 201 |
| 1 | 1024 → 2048 | (8 → 16) × 8 × 8 | 19.0 | 130 | 384 → 768 | (6 → 12) × 4 × 4 | 19.0 | 111 |
| 2 | 2048 → 4096 | 16 × (8 → 16) × 8 | 10.0 | 52 | 768 → 1 536 | 12 × (4 → 8) × 4 | 10.0 | 50 |
| 3 | 4096 → 8192 | 16 × 16 × (8 → 16) | 10.0 | 43 | 1536 → 3072 | 12 × 8 × (4 → 8) | 9.6 | 41 |
| 4 | 8192 → 16384 | (16 → 32) × 16 × 16 | 10.0 | 42 | 3072 → 6144 | (12 → 24) × 8 × 8 | 9.6 | 41 |
| 5 | 16384 → 32768 | 32 × (16 → 32) × 16 | 10.0 | 7 | 6144 → 12288 | 24 × (8 → 16) × 8 | 9.6 | 20 |
| 6 | 32768 → 65536 | 32 × 32 × (16 → 32) | 8.4 | 5 | 12288 → 24576 | 24 × 16 × (8 → 16) | 9.6 | 14 |

**Table 3.** System sizes and integration parameters used in this study. In each case, simulations were performed using the number of atoms indicated as $N_{j+1}$ ($=2N_j$). Integration is performed for $\lambda$ from 0 to $\lambda_m$, using $n_\lambda$ steps, and yields $\Delta A(N_j) \equiv A(N_{j+1}) - A(N_j)$. For hcp, $\lambda$ steps are equally spaced, but in bcc the separation varies along the path.

$$S = \frac{\langle U\rangle - A}{T},\tag{1}$$

which gives us the quantity plotted as the open points in Fig. 1. Here, the average energy is determined from the simulations (with $w = 0$; see Methods section). This average and all other thermodynamic properties computed here are recorded in Table 1 for each system size and crystal structure.

The Gibbs free energy indicates whether hcp or bcc is more stable, with both at 7000 K and 360 GPa. This is obtained via $G(N) = A(N) + PV$, and is computed for each phase as a function of system size and included in Table 1.

## Discussion

Considering Fig. 1, our findings agree with (I) that the entropy favors bcc relative to hcp, but the difference is much smaller than reported for the $N > 20{,}000$ systems in (I), which is taken by the authors of (I) to be where the true bcc behavior is manifested. The discrepancy almost certainly stems from their use of the VACF to estimate the entropy, an approach that must fail when diffusive behavior becomes significant. The weakness of using VACF is acknowledged in (I), which is careful to indicate that the large $S_{bcc}$ represents an upper bound of possible values, while concluding it nonetheless reflects essential behavior that necessitates large-$N$ simulations to compute $S_{bcc}$

rigorously. This conclusion is refuted by our results, which finds little or no variation in $S/N$ with $N$ for the bcc (and hcp) phase, for $N$ as small as 512 and as large as 32,000 atoms.

In addition to the incompatibility of our entropy results and the VACF calculations with respect to variation with $N$, the absolute magnitudes of the entropies from the two methods differ considerably: the VACF values are larger than the entropies that we compute by about 20 J/mol-K. Notably however, the intermediate-$N$ VACF estimate of the bcc-hcp entropy *difference* is in very good agreement with the difference computed from our data. Of course, there is a large region of system sizes where the VACF estimate of the difference is inaccurate, so one must take care not to treat the partial agreement with our results as suggesting that VACF methods are generally reliable in providing entropy differences.

Table 2 records differences in key thermodynamic properties between the phases; these are evaluated directly from the data in Table 1. Enthalpy favors hcp (due mainly to the *PV* contribution), and the entropy difference we compute is not quite sufficient to overcome this to yield a bcc-favorable Gibbs free energy; within uncertainty, they could be in equilibrium at 7000 K and 360 GPa. Undoubtedly then for temperatures $T < 7000$ K the bcc phase for this SC iron model is less stable than hcp, because the enthalpy advantage of hcp becomes even more dominant at lower temperatures. We speculate that evidence in (I) for growth of bcc stability with $N$ based on its persistence in large-system simulations is due to growth of the bcc → hcp activation barrier with $N$, and its reduction with $T$; however, more detailed calculations would be needed to prove this explanation. In any case, rigorous free-energy calculations as an indicator of relative stability must be considered more definitive than conclusions based on the absence of spontaneous transformation.

Our results regarding bcc/hcp stability for the SC model, while rigorous, cannot be considered conclusive with respect to real iron at 7000 K. First, the SC model was fit to DFT energies generated at only one temperature (6000 K), while our analysis is performed at 7000 K to allow direct comparison with SC data reported in (I). Apart from the inaccuracy introduced by not including in the fit the effect of differences in sampling at different states, the effect of temperature on the electronic excitation contribution of the potential is not characterized by the fit as well. In addition, even at 6000 K the SC model exhibits significant differences from accurate *ab initio* DFT calculations for the energy and pressure of crystalline iron, despite being fit to DFT calculations at this temperature. Specifically, isothermal-isobaric simulations we conducted using LAMMPS at 6000 K and 360 GPa with the SC model yielded an average energy ($\langle U \rangle = -5360$ eV (bcc) and $-5490$ eV (hcp) for 1024 atoms) that differs from the range of energies used in (I) to fit the potential (roughly $-5550$ to $-5750$ eV for 1024 atoms, across both bcc and hcp phases). The configurations used in (I) to perform the fit were generated by sampling the DFT potential. Hence, despite its accurate reproduction of the energies of the fitted configurations, this difference in $\langle U \rangle$ suggests that samples generated by the SC model differ from those obtained from the DFT potential it is fit to, even when both are sampled at the same temperature.

Finally, we make a few points related to diffusion in this system. We observed in our simulations the cooperative diffusion events that were reported in (I); these occurred in some systems more than others. A cursory study of the trajectories suggests that they represent the spontaneous generation of a vacancy-interstitial pair that move away from each other after creation. They eventually recombine to stop the process. This recombination occurs more quickly in small systems than in larger ones, which might explain why the phenomenon is observed more in the larger systems. Regardless of the nature of the diffusion events, we find no obvious correlation between the amount of diffusion and the results we obtain for the thermodynamic properties. In fact, we conducted a full study very similar to the one described here, but with a much stronger thermostat (and, incidentally, using a different form for $\phi$ (Eq. (8a)). The stronger thermostat had the (unintended) effect of completely suppressing diffusion. Concerned that this might be perceived as undermining relevance to (I), we repeated all calculations with the weaker thermostat described here, which allows for significant diffusion. The results of the two studies are identical within statistical uncertainty–the presence or absence of diffusion has no effect on the entropy.

In summary, the VACF approach used in (I) fails for systems with diffusion (which is more prevalent for larger $N$) and yielded inaccurate free energies, contributing to the idea that prohibitively large systems are needed to simulate bcc iron properly. Instead, our findings indicate that the bcc entropy per mole is, as expected, nearly constant with system size down to 512 atoms. Consequently it it should be possible to compute accurate free energies of bcc iron without requiring prohibitively large system sizes.

## Methods

**Free-energy calculations.**     The free energy difference $\Delta A(N)$ for a given $N \to 2N$ case is computed via thermodynamic integration. The system is at all points in the integration path composed of $2N$ atoms in a volume of size $2V$. The simulation volume is partitioned conceptually into two equal rectangular subvolumes, each of size $V$; for specificity, we state that one subvolume is translated a distance $L_x$ in the $x$ coordinate direction with respect to the other subvolume, where $L_x$ is the $x$-edge length of each subvolume. Each atom in one subvolume is coupled to a partner atom in the other; label the mutual partners $i$ and $i+N$. Nominally, the partner atoms are at positions where they would be periodic images of each other, and the degree to which this relation is imposed is characterized by the integration parameter $\lambda$. At the start of the integration path, $\lambda = \infty$ and the partner atoms are rigidly coupled, so the configurations sampled within the two subvolumes are identical, and one is truly a periodic image of the other. At the end of the integration path, $\lambda = 0$ and the partner atoms are completely uncoupled. We will use $A_\lambda$ to indicate the $\lambda$-dependent free energy along the path.

The system obtained at the "no-constraint" end of the path is a conventional one having free energy $A(T, 2V, 2N)$; thus

$$A_0(T, \ 2V, \ 2N) = A(T, \ 2V, \ 2N) \tag{2}$$

At the other "rigid" end of the path, the system is closely related to one with $N$ particles in a volume $V$, except that the energy of each configuration is exactly doubled, because it has identical contributions from both subvolumes. To offset this, we double the temperature as well, thus:

$$A_\infty(2T, \ 2V, \ 2N) = A(T, \ V, \ N) \tag{3}$$

Consequently, we require variation of temperature along the path, following $\tau(\lambda)$, such that $\tau(0) = T$ and $\tau(\infty) = 2T$. In terms of the partition function, the free energies are

$$\exp(-A_0/k_B T) = \frac{1}{(2N)!\Lambda(T)^{6N}} Z(\tau(0), 2V, \ 2N; \ \phi(0)) \tag{4a}$$

$$\exp(-A_\infty/k_B T) = \frac{1}{2^N N!\Lambda(T)^{3N}} Z(\tau(\infty), \ 2V, \ 2N; \ \phi(\infty)), \tag{4b}$$

where $\Lambda(T)$ is the de Broglie thermal wavelength, and the $\lambda$-dependent configurational integral is defined:

$$Z(\tau(\lambda), \ 2V, \ 2N; \ \phi(\lambda)) = \int_{2V} \exp\left[-\frac{1}{k_B\tau(\lambda)}\left(U(\mathbf{r}^{2N}) + \sum_{i=1}^{N}\phi(\mathbf{r}_i, \ \mathbf{r}_{i+N}; \ \lambda)\right)\right] d\mathbf{r}^{2N}; \tag{5}$$

the factor of $2^N$ appearing in equation (4b) accounts for the indistinguishability of the partner atoms. Here we have introduced the pair function $\phi(\mathbf{r}_i, \mathbf{r}_{i+N}; \lambda)$, which imposes the partner coupling. Consistent with the integration limits, $\phi(\mathbf{r}_i, \mathbf{r}_{i+N}; 0) \equiv 0$, and $\exp[-\phi(\mathbf{r}_i, \mathbf{r}_{i+N}; \infty)] = \delta(\mathbf{r}_{i+N} - \mathbf{r}_i - \mathbf{L})$, where $\delta$ is the Dirac delta function and $\mathbf{L}$ is the vector $(L_x, 0, 0)$. Conventional periodic boundary conditions are applied to the $(2N, 2V)$ system for all $\lambda$, so all energy functions (including $\phi$) are evaluated using the minimum-image convention.

Combining the above, the working equation for the desired free energy difference is

$$\frac{\Delta A(N)}{k_B T} = \ln\frac{(2N)!}{2^N N!} + 3N\ln\Lambda(T) + \int_0^\infty \frac{d\ln Z}{d\lambda} d\lambda. \tag{6}$$

The integration is performed to a value of $\lambda$ that is large enough to ensure that the partner-atom separation is nearly rigid, meaning that $U$ does not vary significantly for any observed deviations from rigidity. Designating this value $\lambda_m$, equation (6) becomes

$$\frac{\Delta A(N)}{k_B T} = \ln\frac{(2N)!}{2^N N!} + 3N\ln\Lambda(T) + \int_0^{\lambda_m} \frac{d\ln Z}{d\lambda} d\lambda - N\ln\int e^{-\phi(0,r;\lambda_m)/k_B\tau(\lambda_m)} d\mathbf{r}; \tag{7}$$

the $\mathbf{r}$ integral appearing here is the ratio $Z(\tau(\lambda_m), 2V, 2N; \lambda_m)/Z(2T, 2V, 2N, \infty)$ (which assumes $\tau(\lambda_m) \approx 2T$); the integrand diminishes rapidly with $|\mathbf{r}|$, so this integral may be taken over an infinite volume.

A natural choice for the constraint function $\phi$ is a harmonic spring, with width given in terms of $\lambda$:

$$\phi(\mathbf{r}_1, \ \mathbf{r}_2; \ \lambda) = \varepsilon w(\lambda)r^2 \tag{8a}$$

$$r^2 = (|r_{x,1} - r_{x,2}| - L_x)^2 + (r_{y,1} - r_{y,2})^2 + (r_{z,1} - r_{z,2})^2 \tag{8b}$$

where we take $\varepsilon = 10$ J/mol. The rigidity parameter $w$ appearing in equation (8a) is selected as a function of $\lambda$ to support the performance of the free energy calculation. After some trial-and-error, we chose the form satisfying:

$$\lambda = \ln(1 + bw), \tag{9}$$

where $b = e^{-4}$ Å$^2$ is chosen to enhance the smoothness of $d\ln Z/d\lambda$; equation (9) may be easily inverted to obtain $w(\lambda)$.

For $\phi$ given by equation (8a), the r integral in equation (7) is

$$\int e^{-\phi(0,\mathbf{r};\lambda_m)/k_B\tau(\lambda_m)} d\mathbf{r} = \left(\frac{\pi k_B\tau(\lambda_m)}{\varepsilon w(\lambda_m)}\right)^{3/2}. \tag{10}$$

The temperature protocol $\tau(\lambda)$ is constructed using the heuristic of keeping the average energy $\langle U \rangle_\lambda$ roughly constant along the path. This is done also by trial-and-error. We use the following form:

$$\tau(\lambda) = T\left(1 + \frac{w(\lambda)}{[a_0 + a_1 w(\lambda) + w(\lambda)^2]^{1/2}}\right) \tag{11}$$

where $a_0 = 1.2253 \times 10^8$ Å$^{-4}$ and $a_1 = 416305$ Å$^{-2}$. The same protocol $\tau(\lambda)$ was used for all system-size doublings for both bcc and hcp.

Finally, the integrand for the thermodynamic integration in $\lambda$ is given via the ensemble averages (indicated by angle brackets):

$$f(\lambda; N) \equiv \frac{d \ln Z}{d\lambda} - \frac{\overline{U}}{k_B \tau^2}\left(\frac{d\tau}{d\lambda}\right)$$

$$= \left(\left\langle \frac{U - \overline{U}}{k_B \tau^2} \right\rangle_\lambda + \left\langle \frac{\Phi}{k_B \tau^2} \right\rangle_\lambda\right)\left(\frac{d\tau}{d\lambda}\right) + \left\langle -\frac{1}{k_B \tau}\frac{d\Phi}{d\lambda} \right\rangle_\lambda \tag{12}$$

where $\Phi$ is the total constraining potential:

$$\Phi \equiv \sum_{i=1}^{N}\phi(\mathbf{r}_i, \mathbf{r}_{i+N}; \lambda),$$

and $\overline{U}$ is the mean of the energy-averages across the path:

$$\overline{U} \equiv \frac{1}{n_\lambda}\sum_{i=1}^{n_\lambda}\langle U \rangle_{\lambda_i}. \tag{13}$$

We subtract $\overline{U}$ when defining $f$ because the $\tau(\lambda)$ protocol is designed to keep $\langle U \rangle_\lambda$ roughly constant across the integration path. Thus, $\langle U - \overline{U} \rangle_\lambda$ is nearly zero for all $\lambda$ and the first of the three terms on the right-hand side of equation (12) is small. This is good because $(d\tau/d\lambda)$ is not small and has significant curvature, and consequently it would introduce inaccuracy in the quadrature. The bulk of the contribution from integrating $(U/k_B\tau^2)(d\tau/d\lambda)$ is captured by $(\overline{U}/k_B\tau^2)(d\tau/d\lambda)$, for which the integral can be evaluated analytically:

$$\int_0^\infty \frac{\overline{U}}{k_B\tau^2}\frac{d\tau}{d\lambda}d\lambda = \frac{1}{2}\frac{\overline{U}}{k_BT}. \tag{14}$$

We define $\mathcal{J}(N, \lambda_m)$ as the integral of $f$:

$$\mathcal{J}(N, \lambda_m) \equiv \int_0^{\lambda_m} f(\lambda; N)d\lambda. \tag{15}$$

Then in terms of this the free energy difference is:

$$\frac{\Delta A(N)}{k_BT} = \ln\frac{(2N)!}{2^N N!} + \frac{1}{2}\frac{\overline{U}}{k_BT} + \mathcal{J}(N, \lambda_m) + \frac{3}{2}N\ln\left(\frac{\varepsilon w(\lambda_m)\Lambda^2(T)}{\pi k_B\tau(\lambda_m)}\right) \tag{16}$$

The integrand $f(\lambda; N)$ given by equation (12) as computed by molecular simulation is plotted in Fig. 2, parts (a) and (b) for the bcc and hcp systems, respectively. At large $w$ there is no observable difference with system size, but for smaller $w$ the integrands for a given allotrope differ noticeably with $N$. This variation is expected because for larger boxes the atoms have more space to sample, and this factors into the average for smaller $w$, where the finite box volume rather than $\phi$ limits partner separations. The $N$ variation in the integral is compensated by the combinatorial terms involving $N!$ in equation (6), and the overall free energy difference becomes essentially $N$-independent.

The integrands $f(\lambda; N)$ exhibit significant curvature, which introduces inaccuracy in the integrals when quadrature is performed. Given that we are interested in differences in the integrals with respect to $N$, we can offset the problems with curvature by working with differences in the integrands. We find that it is useful to consider up to second differences. Thus, in parts (c) and (d) of Fig. 2 we plot the differences

$$f^{(1)}(\lambda; N) \equiv f(\lambda; 2N) - f(\lambda; N), \tag{17}$$

while in parts (e) and (f) we present

$$f^{(2)}(\lambda; N) \equiv f^{(1)}(\lambda; 2N) - f^{(1)}(\lambda; N). \tag{18}$$

The plots show that these first and second differences are significantly smaller in magnitude than $f(\lambda; N)$ itself and have a much narrower range of $\lambda$ where they are nonzero. The second difference $f^{(2)}$ shows a significant dependence on system size, with some noticeable variation in magnitude and shape for the different $N$. This variation is a result of the variation in the aspect ratio of the box (cubic/oblate/prolate) through successive system doublings. The same behavior was observed in our tests of the method on the fcc Lennard-Jones model, so it does not signify any unusual features of iron or the crystal structures being studied here. The differences average out upon integration over $\lambda$.

We can define a corresponding set of differences in the integrals:

$$\mathcal{J}^{(1)}(N) \equiv \mathcal{J}(2N, \lambda_m) - \mathcal{J}(N, \lambda_m)$$

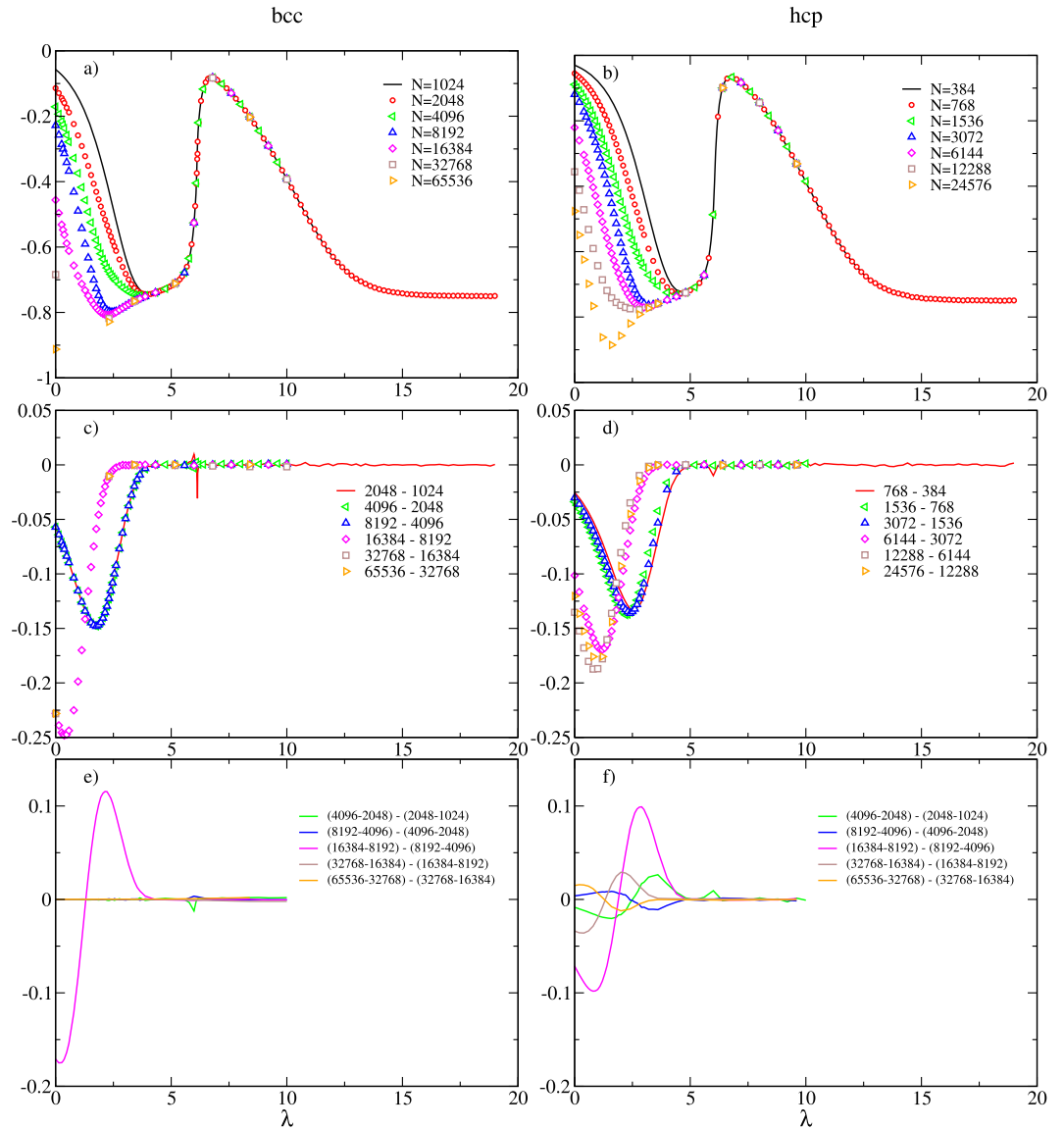$$= \int_0^{\lambda_m} f^{(1)}(\lambda; N)d\lambda \tag{19a}$$

**Figure 2.** Integrand of equation (12), $f(\lambda, N)$ (top row) and its first and second differences $f^{(1)}(\lambda, N)$ (middle row), and $f^{(2)}(\lambda, N)$ (bottom row), respectively. Note that, according to Eq. (20), only the smallest-system $f(\lambda, N)$ data are integrated (black line in top-row plots), and $f^{(1)}(\lambda, N)$ data are integrated for only the smallest pair of systems (red line in middle-row plots). All $f^{(2)}(\lambda, N)$ are integrated. Lines in each figure join points not represented by symbols (no symbols are joined by lines).

$$\begin{aligned}
\mathcal{J}^{(2)}(N) &\equiv \mathcal{J}^{(1)}(2N) - \mathcal{J}^{(1)}(N) \\
&= \int_0^{\lambda_m} f^{(2)}(\lambda; N) d\lambda
\end{aligned} \tag{19b}$$

We suppress indication of dependence on $\lambda_m$ for $\mathcal{J}^{(1)}$ and $\mathcal{J}^{(2)}$ because these are independent of $\lambda_m$ for sufficiently large $\lambda_m$. The first and second integral differences are evaluated using integrals of the $f$ differences, so they can be determined with greater accuracy than the integral $\mathcal{J}(N, \lambda_m)$.

We evaluate $\mathcal{J}(N, \lambda_m)$ directly (i.e., via quadrature according to equation (15)) only for $N_0$, where we have the largest number of quadrature points. Then we determine $\mathcal{J}(N, \lambda_m)$ for larger $N$ by adding the $\mathcal{J}^{(1)}(N_0)$ and $\mathcal{J}^{(2)}(N_j)$ differences, evaluated via numerical integration of $f^{(1)}$ and $f^{(2)}$ according to equation (19). Specifically:

$$\mathcal{J}(N_1, \lambda_m) = \mathcal{J}(N_0, \lambda_m) + \mathcal{J}^{(1)}(N_0) \tag{20a}$$

$$\mathcal{J}(N_2, \lambda_m) = \mathcal{J}(N_1, \lambda_m) + \mathcal{J}^{(1)}(N_0) + \mathcal{J}^{(2)}(N_0) \tag{20b}$$

$$\mathcal{J}(N_n, \lambda_{\mathrm{m}}) = \mathcal{J}(N_{n-1}, \lambda_{\mathrm{m}}) + \mathcal{J}^{(1)}(N_0) + \sum_{k=0}^{n-2} \mathcal{J}^{(2)}(N_k) \qquad (20c)$$

Thus, in Fig. 2 the only data subject to quadrature are those represented by lines joining points. The other data, represented by symbols in the plot, are used only to compute the higher-order differences.

**Simulation details.** We performed canonical-ensemble molecular simulations of the Sutton-Chen potential reported in (I), which was obtained there by fitting to DFT energies of selected configurations of the bcc and hcp phases at 6000 K and 360 GPa; this model truncates interactions at a distance of 6 Å, which puts a lower bound on the system size that may be simulated for a given density. We generated configurations using MD with an Andersen thermostat. All atom velocities were randomized by sampling a Maxwell-Boltzmann distribution every 5000 steps (5 ps).

The relative coordinates of coupled atoms were handled within a multi-timestep integration framework, with the spring force handled analytically in the inner loop. For large values of $w$, the period of the harmonic spring matches the data collection interval such that atoms tend to be at the same relative position each time data is collected, leading to highly correlated results. To break this, we randomized the relative velocity between the atoms in each pair.

For very large values of $w$, the multi-timestep approach fails to properly sample the appropriate distribution of relative coordintes for the coupled pairs. As a remedy, we held $r_{12}$, the vector between atoms 1 and 2, at a fixed value during MD integration and then sampled each component of $r_{12}$ vector from a Gaussian distribution centered at $r_{12} = 0$ with width $\sigma = (k_B \tau / 2\varepsilon w)^{1/2}$. Although the Gaussian distribution perfectly samples the probability due to the spring energy, the iron energy is ignored when selecting new positions. Consequently, we must consider the new positions to be trials that we would accept or reject with Metropolis Monte Carlo (MC) criteria:

$$P_{\mathrm{acc}} = \min(1, e^{-\beta(U_{\mathrm{Fe,new}} - U_{\mathrm{Fe,old}})}) \qquad (21)$$

Simulations were performed for a temperature $T = 7000$ K sampling an NVT ensemble with volumes $V$ selected to result in a pressure of 360 GPa for the given temperature and crystal structure (as given from short NPT runs in LAMMPS): 6.65 Å³/atom for the bcc phase, and 6.616 Å³/atom for hcp. We note that according to (I), the difference in volumes of the two phases at this temperature and pressure when simulated using the DFT potential is about 50% larger than we obtained for the Sutton-Chen potential used here.

Following the scheme outlined in the previous section, system-size doublings were performed to yield free energies from $N = 512$ to 65,536 (seven doublings) for the bcc phase, and $N = 192$ to 24,576 (seven doublings) for the hcp phase (which employs an orthorhombic unit cell with a 4-atom basis). We note that further halving of the smallest systems would result in a simulation volume that is smaller than twice the cutoff radius of the potential, leading to inaccurate results, so these are the smallest systems sizes we can use in this sequence. The shape ratio of the simulation supercell (if unit cell were considered as cubic) cycles from oblate to cubic to prolate and back to oblate in successive doublings. The supercell shapes are recorded along with the system sizes in Table 3. As indicated in the table, we refer to the system sizes from smallest to largest as $N_0, N_1, \ldots, N_7$, such that $N_j = 2^j N_0$.

To obtain correct results, it is necessary for the atoms to explore all positions in the simulation box to the extent allowed by the constraint function $\phi$. Diffusion is not sufficiently fast to ensure this outcome, so additional MC trials were performed in which two atoms selected at random are considered to swap positions; swaps were considered for neighboring atoms within 2 to 6 Å separation (depending on $w$). Acceptance of the trial was made according to the usual Metropolis recipe[7]. These trials require no recalculation of the iron potential energy ($U(r^{2N})$), and involve changes only in the interaction of each atom with its partner via $\phi$, so they are computationally inexpensive and could be performed repeatedly; $10N$ MC swap trials were performed every 10 fs, after completing the MD/MC acceptance decision.

For each system-doubling, simulations were performed for values of $\lambda$ from 0 to $\lambda_{\mathrm{m}}$, and the integrand specified by equation (12) was averaged (separate averages were kept for $U$, $\phi$, and $d\phi/dw$). The specific choice of $\lambda_{\mathrm{m}}$, and number of quadrature points $n_\lambda$ for each $N$-doubling process is recorded in Table 3. To keep computation time manageable, the $n_\lambda$ was reduced about in half for each size doubling (though in some cases additional points were used to ensure the integrand was resolved sufficiently). Each simulation sampled $1 \times 10^5$ MD steps, or 100 ps total; this is after 11.5 ps of equilibration. Hence, the $N_0 \to N_1$ system-doubling involved 100 ps $\times$ 227 $\lambda$ steps, totaling 22.7 ns of sampling, and the $N_6 \to N_7$ case performed a total of about 500 ps of sampling.

Simulation averages for each $\lambda$ were collected into 100 blocks for error analysis; the covariance between successive blocks was examined to ensure they were uncorrelated. Stochastic uncertainties were computed as one standard deviation of the mean (68% confidence level). Inaccuracy (systematic error) in $\Delta A$ arising from the trapezoid-rule numerical integrations was estimated by removing half (or more) of the quadrature points used to compute each integral. In all cases, the inaccuracy is found to be smaller than the stochastic uncertainty.

All calculations reported here were performed using the Etomica molecular simulation package[8]. Tests were conducted of configuration energies to ensure that the energy potential function was consistent with that implemented in LAMMPS, which was used by the authors of (I) in their studies of the Sutton-Chen model. Additionally, we tested the free energy-calculation framework by applying the method to the Lennard-Jones model; the result obtained this way matched accurate values that we obtained independently using more conventional methods[2].

All data generated or analysed during this study are included in this published article.

## References

1. Belonoshko, A. B. *et al*. Stabilization of body-centred cubic iron under inner-core conditions. *Nature Geoscience* **10**, 312–316 (2017).
2. Moustafa, S. G., Schultz, A. J & Kofke, D. A. Harmonically Assisted Methods for Computing the Free Energy of Classical Crystals by Molecular Simulation: A Comparative Study. *J. Chem. Theory Comput.* **13**, 825–834 (2017).
3. Frenkel, D. & Ladd, A. J. C. New Monte Carlo method to compute the free energy of arbitrary solids. application to the fcc and hcp phases of hard spheres. *J. Chem. Phys.* **81**, 3188–3193 (1984).
4. Barnes, C. D. & Kofke, D. A. Self-referential method for calculation of the free energy of crystals by Monte Carlo simulation. *Phys. Rev. E* **65**, 036709 (2002).
5. Sweatman, M. Self-referential Monte Carlo method for calculating the free energy of crystalline solids. *Phys. Rev. E* **72**, 016711 (2005).
6. Sweatman, M. B. Comparison of absolute free energy calculation methods for fluids and solids. *Mol. Phys.* **113**, 1206–1216 (2015).
7. Frenkel, D. & Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications* (Academic Press, San Diego, 2002).
8. Schultz, A. J. & Kofke, D. A. Etomica: An object-oriented framework for molecular simulation. *J. Comput. Chem.* **36**, 573–583 (2015).

## Acknowledgements

## Author Contributions

All authors conceived the calculations, A.S. conducted the simulations, A.S. and D.K. processed the results. All authors contributed to and reviewed the manuscript.

## Additional Information

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.