Partitioning of Sulfur between Solid and Liquid Iron under Earth's Core Conditions: Constraints from Atomistic Simulations with Machine Learning Potentials

Zhigang Zhang[1,2,3,*], Gábor Csányi[4] and Dario Alfè[5,6]

1. Key Laboratory of Earth and Planetary Physics, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, 100029, China
2. State Key Laboratory of Lunar and Planetary Sciences (Macau University of Science and Technology), Macau, China
3. College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
4. Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, United Kingdom
5. Department of Earth Sciences and London Centre for Nanotechnology, University College London, Gower Street, London WC1E 6BT, United Kingdom
6. Dipartimento di Fisica "Ettore Pancini," Università degli Studi di Napoli "Federico II," Complesso di Monte S. Angelo, via Cinthia, 80126 Napoli, Italy

* To whom correspondence should be addressed. E-mail: zgzhang@mail.iggcas.ac.cn, Tel/fax: +86-10-82998622.

ABSTRACT

Partition coefficients of light elements between the solid and liquid iron phases are crucial for uncovering the state and dynamics of the Earth's core. As one of the major light element candidates, sulfur has attracted extensive interests for measuring its partitioning and phase behaviors over the last several decades, but the relevant experimental data under Earth's core conditions are still scarce. In this study, using a toolkit consisting of electronic structure theory, high-accuracy machine learning potentials and rigorous free energy calculations, we establish an efficient and extendible framework for predicting complex phase behaviors of iron alloys under extreme conditions. As a first application of this framework, we predict the partition coefficients of sulfur over wide range of temperatures and pressures (from 4000 K, 150 GPa to 6000 K, 330 GPa), which are demonstrated to be in good agreement with previous experiments and *ab initio* simulations. After a continuous increase below ~250 GPa, the partition coefficient is found to be around 0.75±0.07 at higher pressures and are essentially temperature-independent. Given these predictions, the partitioning of sulfur is confirmed to be insufficient to account for the observed density jump across the Earth's inner core boundary and its roles on the geodynamics and evolution of Earth's core should be minor.

Keywords:

Partition coefficient; Sulfur; Earth's core; First principles; Machine learning; Light elements; Density Functional Theory

1. Introduction

The Earth's core is constituted of some light elements in addition to the major components of iron and nickel (Allegre et al., 1995; Birch, 1964; Mcdonough and Sun, 1995). The prime candidates for these light elements include sulfur, oxygen, silicon, carbon, and hydrogen (Hirose et al., 2013; Li and Fei, 2007; Poirier, 1994). It is recognized that they redistribute across the Earth's inner core boundary (ICB), with more in the outer core (~5-10 wt%) and less in the inner core (~2-3 wt%). The difference in light element contents between the outer and inner core helps explaining the density jump across the ICB (Cao and Romanowicz, 2004; Masters and Gubbins, 2003), anchoring the temperature profile in the whole Earth's core (Morard et al., 2014) and inducing compositional stratification in the inner core (Alboussiere et al., 2010). More importantly, compositional buoyancy created by the redistribution of light elements turns out to be the principal energy source for powering the geodynamo in the outer core (Stacey and Stacey, 1999) and may be critical for driving the convection in the inner core (Gubbins et al., 2013).

The distributions of light elements across ICB are largely controlled by their partition coefficients between the coexisting liquid and solid phases. With several decades' efforts of high-pressure experiments, current knowledge of the partitioning or more broadly phase behaviors of iron alloys has been significantly extended but is still far from adequate (Morard et al., 2014). For sulfur, the focus of this study, the eutectic melting phase relations have been determined by a number of experiments (Chen et al., 2008; Chudinovskikh and Boehler, 2007; Kamada et al., 2012; Kamada et al., 2010; Li et al., 2001; Morard et al., 2014; Morard et al., 2008; Mori et al., 2017; Stewart et al., 2007; Terasaki et al., 2011; Yokoo et al., 2019). Most of these experiments are done at pressures lower than 60 GPa and the data under the real core pressures (>140 GPa) are scarce. Although the highest-pressure record in

these experiments has reached 254 GPa (Mori et al., 2017), it is still some distance away from true ICB condition (330 GPa). With these experimental constraints, people find a general trend of decreasing sulfur contents in the eutectic liquids and increasing solubility of sulfur in the solid solutions at higher pressures, which means an increasing tendency for the partition coefficients of sulfur (Kamada et al., 2012; Morard et al., 2014; Yokoo et al., 2019).

Complementary to high-pressure experiments, first principles simulations (mostly based on Density Functional Theory (DFT)) provide an alternative route to unravel the partitioning and phase behaviors of iron alloys under extreme conditions. While early DFT simulations have already predicted the partition coefficient of sulfur under ICB conditions (Alfe et al., 2002a, 2003, 2007), further work is needed to establish consistency with the experimentally measured values due to the gap in the *T-P* regime accessible to the experiments and first principles simulations. The reason for the lack of extensive first principles simulations for the binary or more complex iron alloying systems is the extremely high computational cost of such simulations, typically taking many orders of magnitude more than corresponding simulations employing empirical atomistic potentials, which however are generally not accurate enough to give meaningful predictions at extreme conditions.

In this study, we are taking a different approach and use a new generation of *surrogate models* in place of DFT, interatomic potentials made using machine learning techniques (Rupp, 2015). The basic idea of this approach is to get a relatively small amount of DFT data (total energies, forces and stresses from a short MD trajectory) and construct nonparametric potentials that approximate the true *ab initio* Born-Oppenheimer potential energy surface very closely and then carry out the extensive sampling using the potentials. With newly constructed highly accurate potentials for the Fe-S binary system, combining with rigorous free energy calculations, our efforts in this study not only extend the partition

coefficients of sulfur to the overlapping *T*-*P* regime of experiments and simulations but also provide a general and extendible framework for effectively predicting the phase behaviors of multi-component iron alloying systems at extreme conditions.

2. Methods

2.1. Machine learning potentials

We use the Gaussian Approximation Potential (GAP) (Bartok et al., 2010), essentially a kernel ridge regression method (Kung, 2014). This is just one of a class of recently popularized machine learning methods for creating nonparametric interatomic potentials, which has been shown to be very successful in tackling difficult materials modelling problems, ranging from investigating the structure of amorphous materials (carbon (Deringer et al., 2017, 2019), silicon (Bartók et al., 2018)), the mechanics of metals (tungsten (Szlachta et al., 2014), iron (Dragoni et al., 2018)) to molecular liquids such (water (Bartók et al., 2013a), methane (Veit et al., 2019). There are many alternatives, using other regression frameworks, such as artificial neural networks (Behler and Parrinello, 2007) and even linear regression (Shapeev, 2017; Drautz, 2019). All these methods are improvable, since using more input data typically leads to more accurate potentials, due to the nonparametric nature of the functional forms.

The theoretical details of the GAP model can be found elsewhere (Bartók et al., 2013b; Ceriotti et al., 2018), we only give a high-level description of the key formula here. Assuming the Born-Oppenheimer potential energy surface of a set of atoms is a smooth function of the atomic coordinates, we write the total energy as a sum of atomic contributions

$$E = \sum_i \varepsilon(\boldsymbol{q}_i) \qquad (1)$$

where the short-ranged local atomic energy ε is assumed to depend explicitly on the positions of the

atoms within a sphere of radius r_{cut} centered on atom i and $\boldsymbol{q}_i$ is a vector representing its local atomic environment. In the GAP framework, Gaussian process regression is used to model ε,

$$\varepsilon(\boldsymbol{q}) = \sum_s \alpha_s K(\boldsymbol{q}_s, \boldsymbol{q}) \equiv \mathbf{k}(\boldsymbol{q})^{\mathrm{T}}\boldsymbol{\alpha} \qquad (2)$$

Given two atoms i and j, the kernel function $K(\boldsymbol{q}_i, \boldsymbol{q}_j)$ is the expected covariance of the respective local atomic energies $\varepsilon(\boldsymbol{q}_i)$ and $\varepsilon(\boldsymbol{q}_j)$, and can be interpreted as a measure of similarity of the two local atomic environments. We choose the Smooth Overlap of Atomic Positions (SOAP) kernel (Bartók et al., 2013b),

$$K(\boldsymbol{q}_i, \boldsymbol{q}_j) = \sigma_w^2 |\widehat{\boldsymbol{q}_i} \cdot \widehat{\boldsymbol{q}_j}|^{\xi} \qquad (3)$$

where the descriptor $\widehat{\boldsymbol{q}_i}$ is the rotational power spectrum of the atomic neighbor density, which is a smooth and regular function, invariant to rotation and permutation of like atoms. The key advantage of the SOAP representation is that there are very few empirical parameters needed, basically the cutoff distance, the Gaussian smearing of atomic positions (essentially band-limiting the spherical Fourier transform of the neighbor density), and small integer power ξ to which the linear SOAP kernel is raised to, controlling the body order of the resulting force field, such that higher exponents result in higher order many-body terms. All the hyper-parameters, including those inherent in the definition of the rotational power spectrum are listed in Table 1.

2.2. Thermodynamics for predicting the partition coefficient

For the Fe-S binary systems, the partition coefficient of sulfur and the other melting properties are determined by the following chemical equilibrium between the liquid and solid phase:

$$\mu_{\mathrm{S}}^{\mathrm{liq}}\left(T, P, c_{\mathrm{S}}^{\mathrm{liq}}\right) = \mu_{\mathrm{S}}^{\mathrm{sol}}(T, P, c_{\mathrm{S}}^{\mathrm{sol}}) \qquad (4)$$

$$\mu_{\mathrm{Fe}}^{\mathrm{liq}}\left(T, P, c_{\mathrm{S}}^{\mathrm{liq}}\right) = \mu_{\mathrm{Fe}}^{\mathrm{sol}}(T, P, c_{\mathrm{S}}^{\mathrm{sol}}) \qquad (5)$$

where μ_{S} and μ_{Fe} are the chemical potential of sulfur and iron respectively, the superscripts liq and sol denote liquid and solid phase respectively and c_{S} is the mole fraction of sulfur in the solution.

Since the chemical potential μ_{S} diverges logarithmically in the low-concentration limit ($c_{\mathrm{S}}\rightarrow 0$), at each temperature and pressure it is useful to express μ_{S} as (Alfe et al., 2002a)

$$\mu_{\mathrm{S}} = k_B T ln c_{\mathrm{S}} + \overline{\mu_{\mathrm{S}}}(c_{\mathrm{S}}) \quad (6)$$

where $\overline{\mu_{\mathrm{S}}}(c_{\mathrm{S}})$ is well behaved for all concentration. While many models have been proposed for $\overline{\mu_{\mathrm{S}}}(c_{\mathrm{S}})$ (or its equivalences), such as various symmetric/asymmetric regular solution models (White, 2013) or those proposed by Ma (2001), the following expansion for $\overline{\mu_{\mathrm{S}}}(c_{\mathrm{S}})$ is simple but practically meaningful

$$\overline{\mu_{\mathrm{S}}}(c_{\mathrm{S}}) = \mu_{\mathrm{S}}^{\dagger} + \lambda_{\mathrm{S}} c_{\mathrm{S}} + O(c_{\mathrm{S}}^2) \quad (7)$$

In line with previous studies (Alfe et al., 2007; Gubbins et al., 2013; Labrosse, 2014), we neglect higher order terms ($O(c_{\mathrm{S}}^2)$) in this study. Therefore, only two parameters of $\mu_{\mathrm{S}}^{\dagger}$ and λ_{S} are involved in the calculations and we obtain the following equation for the chemical potential of sulfur

$$\mu_{\mathrm{S}} = k_B T ln c_{\mathrm{S}} + \mu_{\mathrm{S}}^{\dagger} + \lambda_{\mathrm{S}} c_{\mathrm{S}} \quad (8)$$

According to the Gibbs-Duhem equation, we can straightforwardly get the chemical potential of iron

$$\mu_{\mathrm{Fe}} = \mu_{\mathrm{Fe}}^{0} + (k_B T + \lambda_{\mathrm{S}}) ln(1 - c_{\mathrm{S}}) + \lambda_{\mathrm{S}} c_{\mathrm{S}} \quad (9)$$

where μ_{Fe}^{0} is the chemical potential of pure iron at the same temperature and pressure, which is the Gibbs free energy per atom of pure iron.

Then the Gibbs free energy (G) of the whole system can be expressed by

$$G = N_{\mathrm{S}}\mu_{\mathrm{S}} + N_{\mathrm{Fe}}\mu_{\mathrm{Fe}} = N\left[c_{\mathrm{S}}\mu_{\mathrm{S}} + (1 - c_{\mathrm{S}})\mu_{\mathrm{Fe}}\right] = N\{[c_{\mathrm{S}} ln c_{\mathrm{S}} + (1 - c_{\mathrm{S}})\ln(1 - c_{\mathrm{S}})]k_B T + (1 - c_{\mathrm{S}})\mu_{\mathrm{Fe}}^{0} + c_{\mathrm{S}}\mu_{\mathrm{S}}^{\dagger} + [c_{\mathrm{S}} + (1 - c_{\mathrm{S}})\ln(1 - c_{\mathrm{S}})]\lambda_{\mathrm{S}}\} \quad (10)$$

where N_{Fe} is the number of iron atoms, N_S is that of sulfur atoms in the system and $N = N_{Fe} + N_S$.

Now with the simulated free energies of both liquid and solid phases at several concentrations, we can regress the effective values of $\mu_S^\dagger$ and λ_S with Eqn. (10) and the chemical potentials of sulfur and iron can be calculated thereafter with Eqns. (8) and (9). Finally, the partition coefficient of sulfur (D_S) can be derived from the equilibrium compositions from Eqns. (4) and (5), i.e., $D_S = c_S^{sol}/c_S^{liq}$.

2.3. Free energy calculations

In this study, we calculated the Helmholtz free energies (F) using thermodynamic integration, which rigorously relate the free energy (F_1) of the target system with that of a reference system (F_0) by

$$F_1 = F_0 + \int_0^1 \frac{\partial F_\lambda}{\partial \lambda} d\lambda \quad (11)$$

where λ is a coupling parameter relates the two systems and F_λ is the free energy of the system with total energy U_λ, which is a hybrid total energy potential with the property of being equal to the reference total energy U_0 for λ=0 and the target total energy U_1 for λ=1. If we choose $U_\lambda=(1-\lambda)U_0+\lambda U_1$, then the integrand in Eqn. (11) becomes the energy difference as a function of λ, $\frac{\partial F_\lambda}{\partial \lambda} = \langle U_1 - U_0 \rangle_\lambda \equiv \overline{\Delta U}(\lambda)$, where $\langle \quad \rangle_\lambda$ means thermal average in the ensemble generated by U_λ.

2.3.1. Liquid phases

For the liquids, we choose the ideal gas at the same temperature and composition as the reference system, i.e.,

$$F_0^{ideal} = -k_B T \ln\left(\frac{V^{N_{Fe}+N_S}}{N_{Fe}!N_S!\Lambda_{Fe}^{3N_{Fe}}\Lambda_S^{3N_S}}\right) \quad (12)$$

where $\Lambda=h/(2\pi M k_B T)^{1/2}$ (h is the Planck constant, k_B is the Boltzmann constant, M is the atomic mass

of iron or sulfur and T is the temperature) is the thermal wavelength of iron or sulfur. Note that in standard statistical mechanics the Stirling approximation is usually written as $\ln N! \approx N\ln N - N$, but since we are dealing with small values of N, the natural logarithm of first-order Stirling's series $\ln N! \approx N\ln N - N + \frac{1}{2}\ln(2\pi N)$ was used instead in the calculation of Eqn. (12) to get more accurate free energy (Arfken, 1985).

To integrate from ideal gas to the target system, since the energetic changes are large, especially near the end-points, we used the Gauss-Lobatto quadrature by the following variable transformation

$$\lambda(x) = \left(\frac{x+1}{2}\right)^{\frac{1}{1-k}} \quad (13)$$

with x spans from -1 to 1. Then the integral in Eqn. (11) becomes

$$\int_0^1 \frac{\partial F_\lambda}{\partial \lambda} d\lambda = \frac{1}{2(1-k)} \int_{-1}^{1} \overline{\Delta U}(\lambda(x))\lambda(x)^k dx \approx \frac{1}{2(1-k)} \sum_{i=0}^{n} \omega_{x_i} \overline{\Delta U}(\lambda(x_i))\lambda(x_i)^k \quad (14)$$

k in Eqns. (13) and (14) is selected to be 0.8, which avoids the endpoint singularity since the $\overline{\Delta U}\big(\lambda(x_i)\big)\lambda(x_i)^k$ can be safely set to be zero when λ=0 without loss of accuracy (Dorner et al., 2018). In Fig. A.1(a) of Appendix A, we show a typical transformed integrand as a function of the integration variable x.

The weight functions ω_{x_i} in Eqn. (14) can be calculated by

$$\omega_{x_i} = \frac{2}{n(n-1)[P_{n-1}(x_i)]^2} \quad (15)$$

where P_n are the Legendre polynomials. Practically we include 8 abscissas (n=8) in our calculations of Eqn. (14).

2.3.2. Solid phases

For the pure iron systems, we choose the harmonic *hcp*-lattices at the same temperature and composition as the reference systems, whose free energies can be calculated through lattice dynamics

197 and quasi-harmonic approximation theory (Alfè, 2009). From the harmonic system to the target system,
198 since the energetic changes are small and vary smoothly with the integration variable (as shown in Fig.
199 A.1(b) of the Appendix A), we use the simpler three-point Gauss-Legendre quadrature by the following
200 linear transformation

201 $$\lambda(x) = \frac{x+1}{2} \quad (16)$$

202 with *x* spans from -1 to 1. Then the integral in Eqn. (11) becomes

203 $$\int_0^1 \frac{\partial F_\lambda}{\partial \lambda} d\lambda = \frac{1}{2}\int_{-1}^{1} \overline{\Delta U}(\lambda(x))dx \approx \frac{1}{2}\sum_{i=1}^{3} \omega_{x_i}\overline{\Delta U}(\lambda(x_i)) \qquad (17)$$

204 For a solid solution of Fe-S alloy under specific *T-P* condition, we swapped the sulfur atoms into
205 irons and calculated the Helmholtz free energy change at the same volume through thermodynamic
206 integration (Eqn. (11)). In this case, the energy barriers for interchanging the atoms at various lattice
207 sites greatly hinder a full sampling of the phase space through simple molecular dynamic simulations
208 and the results would inevitably depend on the initial configurations. To circumvent this problem and
209 to avoid exhaustively sampling all the possible configurations with distinct sulfur orderings, which is
210 infeasibe even considering the reducible symmetries, we used the more efficient hybrid Monte
211 Carlo/molecular dynamics simulations (Widom et al., 2014). Since the swapping of species changes
212 the identities and masses of the atoms, the variations of ideal lattice gas contributions to the free
213 energies, which can be calculated through Eqn. (12), need to be counted in the free energy changes.
214

215 **3. Simulation details**

216 3.1. First principles simulations

217 To train the GAP models, we carried out a number of density functional theory (DFT) simulations
218 with VASP, with the projector-augmented-wave method (Kresse and Joubert, 1999). We use the PBE

219 form of Generalized Gradient Approximation (GGA) (Perdew et al., 1996) with valence electrons of
220 16 (valence configuration $3s^2 3p^6 3d^7 4s^1$) for iron, which has been demonstrated to closely resemble the
221 all-electron potential and be important for obtaining accurate melting properties (Sun et al., 2018), and
222 6 ($3s^2 3p^4$) for sulfur.

223 Extensive molecular dynamics (MD) simulations have been deployed to sufficiently sample the
224 *T-P-X* space spanning over 0-25 at. % sulfur concentrations, from 1000 K to 7000 K and from about
225 50 GPa to 500 GPa for the solid phases, and from 4000 K to 10000 K and from about 100 GPa to 500
226 GPa for the liquid phases. Overall 243 simulations have been carried out for the solid phases and 215
227 simulations for the liquid phases. The solid solutions were initiated with hexagonal close-packed (*hcp*)
228 structures with 96 atoms in the supercells (4×4×3 extension of the unit-cell). At high enough
229 temperatures, the solids would always be melted and the final configurations of these runs were
230 adopted to initialize liquid simulations. MD trajectories were propagated in the *NVT* ensemble with
231 the Nosé thermostat for 3~6 ps. At each time step (1 fs interval), the electronic structure is calculated
232 at the Brillouin zone center with an energy cutoff (ENCUT) of 500 eV and iteration convergence
233 criterion (EDIFF) of 10^{-6} eV.

234 From these sampled configurations, we extracted independent configurations every 100-150 steps
235 and re-calculated their energies and forces with much higher precisions. The energy cutoff was
236 increased to 800 eV and the Brillouin zone sampling grid of spacing (KSPACING) was set to be 0.3
237 $Å^{-1}$, which is well converged as revealed by our benchmarks with finer spacing of 0.2 $Å^{-1}$. The
238 projection operators were evaluated in reciprocal space (LREAL=.FALSE.). As in the MD simulations,
239 the influences of finite temperature on the electronic free energy were accounted through the Mermin
240 functional (Mermin, 1965) with Fermi-Dirac smearing of electron occupancy (ISMEAR= −1) and an

241 electronic temperature equal to that of the ions.

242

243 3.2. Training of GAP models

244 The GAP models were trained through the QUIP code (Bartok et al., 2010) based on the high
245 precision DFT simulation results. Since the potential energy surfaces sampled by these simulations are
246 implicitly a function of the electronic temperature, at each temperature we obtained DFT energies
247 that include electronic-entropy, the forces and stresses from over 1000 configurations and organized
248 them into an extended xyz file for the GAP model training, as explained in Appendix B.

249 We list the main hyper-parameters of trainings for the GAP models in Table 1. In these
250 parameters, the zero points of the energy per atom (e0) of iron and sulfur find their contributions mostly
251 from the electronic entropy and therefore are temperature-dependent. They were evaluated by single-
252 step DFT calculations with only one atom in vacuum (without the effects of the periodic images).
253 Finally, to keep models robust at very short interatomic distances, which may not be well sampled in
254 the DFT-MD simulations, we explicitly calculated the dimer potentials for pairs of Fe-Fe, Fe-S and S-
255 S and used their repulsive part as baselines. Details for a typical GAP model training with the QUIP
256 code can be found in Appendix B.

257

258 3.3. Atomistic simulations with the GAP models

259 With the derived GAP models, we carried out three types of atomistic simulations in this study:
260 simple molecular dynamics simulations for pure iron and Fe-S liquids, hybrid Monte Carlo/molecular
261 dynamics simulations for Fe-S solid solutions, and lattice dynamics simulations for pure solid iron to
262 get the reference harmonic free energy. The energetics and forces in all these simulations were

263 calculated with the engine of VASP by invoking a custom-made interface to QUIP. External scripts
264 were utilized to realize the construction of Metropolis Markov chain in Monte Carlo move and relevant
265 lattice dynamics calculations.

266 For simple molecular dynamics simulations, at each T and composition X, we carried out an NPT
267 simulation to evaluate the volume and lattice parameters at specific pressure. Parrinello-Rahman
268 dynamics with Langevin thermostat was adopted to control the pressure and temperature in the
269 trajectory. For the liquid phases, the unit cells were constrained to be cubic throughout the simulations;
270 while for the solid iron, the lattice parameters are constrained to be orthorhombic with b/a ratio fixed
271 to that of the *hcp*-lattice. In both the liquid and solid phases, we generally used 180 atoms, which is
272 demonstrated to be large enough to get converged equations of state and the free energy changes with
273 respected to the referenced systems (Sun et al., 2018). By discarding the first 5 ps (5000 MD steps) as
274 pre-equilibrium stage, we obtained the averaged lattice parameters from the last 20 ps (20000 MD
275 steps) trajectories. Then we carried out an NVT simulation with these lattice parameters, confirming
276 that the system was in a hydrostatic state with deviatoric stresses less than 0.5 GPa.

277 For the solid solutions, at each temperature and composition, with sulfur atoms initially randomly
278 substituted on the *hcp*-Fe lattice sites, we carried out the hybrid Monte Carlo/molecular dynamics
279 simulations (Widom et al., 2014) in a sequence of every 20 MD steps followed by one attempted Monte
280 Carlo (MC) species swap. With a duration of overall 20 ps MD steps and 1000 attempted species swaps,
281 the simulation finds its convergence with well-sampled pressures/volumes and energies and this turns
282 out to be important for the free energy calculations.

283 Finally, lattice dynamics simulations were deployed with PHON through small displacement
284 method (Alfè, 2009). Based on the final configurations of above-mentioned equilibrated molecular

dynamics simulations, with relaxations of the time-averaged ionic positions, the interatomic force constants were calculated by setting a displacement amplitude of 0.01 Å for each atom. A dense 30×30×30 mesh was used for the q-point sampling in the first Brillouin zone to evaluate the vibrational density of state. For each structure of solid phase, we carefully confirmed that the phonon spectra are dynamically stable with no imaginary phonon frequencies, as demonstrated in Fig. A.2 of Appendix A. In the framework of quasi-harmonic approximation theory, the harmonic free energies at finite temperatures can be calculated.

4. Results

4.1. Benchmarks of the GAP models

As the fundamental benchmarks of the derived GAP models in this study, we systematically compare their atomic forces, stresses and energies with those from DFT simulations. As shown in Fig. 1(a), which includes 344928 data retrieved from both simulations at 6000 K, we find very good agreements between the GAP forces and DFT forces. The deviations of the atomic forces are generally within ±2 eV/Å and the average error is around 0.2 eV/Å. The GAP model also reproduces the stress tensors very well, with averaged error of the normal components less than 0.8 GPa as illustrated in Fig. 1(b). Furthermore, it is remarkable that the GAP energies agree excellently with those from DFT simulations, with most of the deviations in GAP energy within ±10 meV/atom and the averaged energy error of 3.6 meV/atom as shown in Fig. 1(c). These accurate depictions of microscopic interactions and reproductions of the energies guarantee the almost identical samplings of the phase spaces with the GAP models as compared with those of the DFT simulations.

To further demonstrate the accuracy of the GAP models for the free energies, we carried out a

benchmark simulation with the GAP model at 6000 K. The *NVT* simulation trajectory is propagated in a 4.6 cm^3/mol cell with 88 iron atoms and 8 sulfur atoms (corresponding to about 237 GPa). We randomly picked out 20 independent configurations from the trajectory and carried out high precision DFT simulations (with parameters mentioned in the second part of Sec. 3.1). Since the energies are sufficiently close between GAP and DFT, through the one-step thermodynamic perturbation method, $\Delta F = k_{\mathrm{B}} T ln\langle e^{-\Delta U/k_{\mathrm{B}}T}\rangle$, we estimated the free energy deviation to be only about 0.3 meV/atom (0.029 kJ/mol).

In Fig. 2, we show the benchmark of the GAP models for the microscopic structures. It is evident that the simulation with GAP models has perfectly sampled the local structures around different atoms as compared with those from the DFT simulation. Just as previous findings from first principles simulations (Alfe and Gillan, 1998; Alfe et al., 2003), we can easily observe the net S-S repulsions and Fe-S attractions in the solution. And almost all the structural features (positions of local maximums and minimums, and spreading of the peaks) of the Fe-S system have been reproduced in the GAP model simulation. Since the radial distribution functions are closely related with the free energy (Ben-Naim, 2006), this benchmark again verifies the robustness of the GAP models in predicting the thermodynamic properties of Fe-S systems.

4.2. Free energy and melting of pure iron

To determine the melting points of pure iron, we firstly carried out some explorative simulations for its equation of states. As shown in Fig. 3(a), squeezing the liquid or expanding the solid inevitably find phase transitions when the pressure deviates far enough from the melting point (P_{m} with uncertainty in the figure, determined below). We carefully inspect the radial distribution functions,

329 mean square displacements and fluctuations of energies and pressures in the trajectories to verify the
330 stabilities of the simulations within the simulation length scales. Based on these points we get the
331 auxiliary Birch-Murnaghan equation of state for the liquid and solid respectively, which are shown in
332 Fig. 3(a) with dashed curves and the numerical data can be found in Appendix C. It is apparent that
333 the melted solids quickly establish new equilibrium within the duration of the simulation and their *P*-
334 *V* relations match the predictions of the liquid EOS quite well, while the solidifying liquids take much
335 longer time to be fully crystalized and the averaged volumes/pressures are still slightly larger than the
336 solid EOS predictions. These points give us rough estimates of the upper and lower limits of the
337 melting pressure. In Fig. 3(a), we also include the curves provided by Komabayashi and Fei (2010)
338 based on existing experimental data and find very good agreements with the simulated EOS.

339 Based on the observations of these simulations, for each temperature we chose a point to calculate
340 its free energy directly with the thermodynamic integration techniques. As listed in Table 2, we got the
341 free energies of eight points, which were used as references for the profiles over the whole pressure
342 range with the aid of the auxiliary equation of state mentioned above. Then through the crossover of
343 these Gibbs free energy profiles of the liquid and solid phases we finally determined the melting points,
344 as demonstrated in Fig. 3(b). The uncertainty of melting pressure was estimated as the half of
345 distance between the upper and lower bound of the crossover by counting the errors of free energies.

346 Fig. 4 shows the calculated melting points of iron at four temperatures. By interpolating these
347 points with simple second-order polynomial, we obtained the following equation for the simulated
348 melting temperature (T_m) of pure iron (valid from about 110 GPa to 430 GPa)

349 $$T_m = 2572 + 14.094 \times P - 0.00891 \times P^2 \quad (18)$$

350 From this equation and considering the error bars of the four points in Fig. 4, the melting temperature

351 at 330 GPa is estimated to be 6253±170 K, comparing very well with 6170±200 K as recently estimated
352 by Sun et al. (2018) and 6350±300 K as predicted by Alfe et al. (2002c). In contrast, simulations by
353 Belonoshko et al. (2000) and Laio et al. (2000) give significantly higher (7100 K) or lower (5400 K)
354 melting temperature, respectively, at the inner core boundary. With careful evaluations of the free
355 energy errors from the classical potentials used in their studies, it is possible to correct these results to
356 be in much better accordance with the DFT simulations (Alfe et al., 2002b).

357 Compared with the experiments, it is remarkable that our results almost perfectly agree with the
358 measurements and extrapolations by Anzellini et al. (2013) over the entire pressure range. Shock wave
359 measurements by Nguyen and Holmes (2004) and Brown and McQueen (1986) also fall into the same
360 trend. On the other hand, diamond anvil measurements by Boehler (1993) and Sinmyo et al (2019) and
361 shock wave measurements by Yoo et al. (1993) give much lower or higher melting temperatures by up
362 to over 1000 K. The exact reasons for these remarkable observed differences in experiments are still
363 under debate due to the extreme technical challenges (Aquilanti et al., 2015; Morard et al., 2018).

364

365 4.3. Free energies of Fe-S alloys

366 For the Fe-S alloys, we deployed simulations with 9-36 sulfur atoms in the 180-atom cells
367 (corresponding to c_S=0.05~0.20, which is within the stability regime of *hcp*-structured Fe-S solid
368 solutions according to previous studies (Cote et al., 2008; Gavryushkin et al., 2016)) and for each
369 concentration we calculated its free energies through the thermodynamic integration techniques at
370 several *T-P* conditions (as listed in Table 3), based on which the free energies over the pressures can
371 be derived through the auxiliary Birch-Murnaghan equation of states at various temperatures, as listed
372 in Appendix C.

In Fig. 5, we show the simulated Gibbs free energies of Fe-S at 4000 K, 250 GPa and 6000 K, 330 GPa. We include the free energies of solid solutions calculated from thermodynamic integrations based on simple MD simulations (open blue circles) to emphasize the importance of enhanced sampling. With quasi-random initial occupations of sulfur atoms on the *hcp*-lattice sites, these free energies agree with those from more sophisticated hybrid MC/MD samplings at low concentrations, but the deviations become more and more severe at high concentrations. The clues for explaining such errors can be found in Fig. 6, which illustratively compares the evolutions of thermodynamic properties and structures in the simulations at 4000 K, 4.3893 cm^3/mol and c_S=0.20. From Fig. 6(a), obviously the hybrid MC/MD sampling quickly finds a more stable state with total energy decreased by over 8.4 kJ/mol within about 4000 steps (involving only 200 attempted MC swapping of the species). The major structural feature of this more stable state is the decreased number of nearest-neighbor S-S pairs ($NN_{S\text{-}S}$), as shown in the bottom plot of Fig. 6(a) and demonstrated in Fig. 6(b) by the radial distribution functions. $NN_{S\text{-}S}$ decreases from the initial quasi-random arrangement value (around 43 here for a 180-atom cell with 36 sulfur atoms, i.e., $NN_{S\text{-}S}=0.5zN_Sc_S$, where z=12 for the *hcp*-lattice) to less than half of the initial value (around 19 in Fig. 6), and the S-S coordination number decreases from 2.5 to 1.1. Since the interactions among nearest-neighbored S-S pairs dominate the energetic change of the Fe-S solution as compared with pure iron (Alfe et al., 2002a), the net repulsions of S-S interactions here sampled by the hybrid Monte Carlo/molecular dynamics simulations effectively push the systems to lower energy states. As shown in the middle plot of Fig. 6(a), the accommodations of sulfur atoms in the more appropriate sites result in a smaller pressure by over 4.2 GPa, which would further contribute to the decreasing of free energy. For the readers' reference, the converged configurations of Fe-S solid solutions through hybrid Monte Carlo/molecular dynamics simulations at various conditions can be

found in Appendix C.

4.4. Partitioning of sulfur

From the data points in Fig. 5, we obtained the free energy profiles over the composition range by linear least-square regressions with Eqn. (10). The partitioning of sulfur in the coexisting liquid and solid iron phases can be firstly inferred from the variations of the two regressed parameters over temperatures and pressures. As listed in Table 4, the difference of $\mu_{\mathrm{S}}^{\dagger,\mathrm{sol}} - \mu_{\mathrm{S}}^{\dagger,\mathrm{liq}}$ is positive. It is almost invariant to pressure but shows clear temperature dependence: it is about 65 kJ/mol at 4000 K, around 21 kJ/mol at 5000 K and 13 kJ/mol at 6000 K. These observations demonstrate that sulfur would generally prefer the liquid iron phase, as expected, but this tendency would be decreased at higher temperatures. On the other hand, the difference of $\lambda_{\mathrm{S}}^{\mathrm{sol}} - \lambda_{\mathrm{S}}^{\mathrm{liq}}$ keeps increasing from -192 kJ/mol at 4000 K and 150 GPa to 135 kJ/mol at 6000 K and 330 GPa. This opposite trend would greatly balance the partitioning of sulfur between the two phases.

By solving Eqns. (4-5), we quantitatively calculated the partition coefficients of sulfur ($D_{\mathrm{S}} = c_{\mathrm{S}}^{\mathrm{sol}}/c_{\mathrm{S}}^{\mathrm{liq}}$), as shown in Fig. 7. At 250 GPa and 330 GPa, similar values of D_{S} are predicted over temperature change of 1000 K, which reveals the minor temperature dependence of D_{S} and this is in accordance with the experimental findings (Kamada et al., 2010; Yokoo et al., 2019). For pressures lower than 250 GPa, we find remarkable pressure dependence of D_{S}: its value decreases by over 40% from 250 GPa to 150 GPa. The simulated data of D_{S} in this study perfectly match the available experimental measurements and their trends up to 254 GPa (Kamada et al., 2012; Kamada et al., 2010; Li et al., 2001; Mori et al., 2017; Stewart et al., 2007; Yokoo et al., 2019). From 250 GPa to higher pressures, it is notable that D_{S} becomes almost pressure independent with an averaged value of

417 0.75±0.07. This is in good agreement with the early prediction by Alfe et al. (2002a) through DFT
418 simulations. Finally, the simple MD simulations with random solid solutions turn out to significantly
419 underestimate D_S (~0.53±0.04 as illustrated in Fig. 7), which again emphasizes the importance of
420 sufficient sampling with the hybrid Monte Carlo/molecular dynamics simulations.
421

422 **5. Implications**

423 According to our simulations in this study, the melting of pure iron at Earth's current inner core
424 boundary pressure (330 GPa) results in a density jump ($\Delta\rho \equiv \rho^{sol} - \rho^{liq}$) of 0.24 g/cm^3, which amounts to
425 1.8% of ρ^{sol} (i.e., $\Delta\rho/\rho^{sol} \approx 1.8\%$) and compares very well with those predicted by previous simulations
426 (as listed in the Table II of Sun et al. (2018)). This density jump is far smaller than the seismologically
427 observed value of 0.6 g/cm^3 (Dziewonski and Anderson, 1981) or 0.8±0.2 g/cm^3 (Masters and Gubbins,
428 2003). Assuming Fe-S binary model for the Earth's core and the temperature around 6000 K, we would
429 need around 20% sulfur ($c_S^{liq} \approx 0.20$) to match the outer core density (ρ^{sol}=12.166 g/cm^3, according to
430 Dziewonski and Anderson (1981)) at ICB. With smaller amount of sulfur coexisting in the inner core,
431 $c_S^{sol} = D_S \times c_S^{liq} = 0.15$, the density jump $\Delta\rho$ increases to 0.39 g/cm^3. So the partitioning with sulfur
432 alone is insufficient to account for the observed density jump across the Earth's inner core boundary
433 and the Earth's core composition should be more complex than the simple Fe-S binary, which supports
434 the findings of previous studies (Alfe et al., 2002a, 2003, 2007). Although the exact core compositions
435 are still elusive, geochemical studies often place around 2 wt% sulfur in the Earth's core (Allegre et
436 al., 1995; McDonough, 2003; Wood et al., 2006). With this amount of sulfur, according to the
437 comprehensive explorations by Badro et al (2014), the outer core would additionally need about 2.6
438 wt% oxygen and 1.8 wt% silicon to best fit the densities and seismic velocities at ICB and CMB (core

439 mantle boundary). This leads to an outer core composition with $c_{\mathrm{O}}^{\mathrm{liq}} = 0.083$, $c_{\mathrm{Si}}^{\mathrm{liq}} = 0.033$ and
440 $c_{\mathrm{S}}^{\mathrm{liq}} = 0.032$. Since D_{Si}≈1, D_{O}<0.01 (calculated from the parameters provided by Alfe et al (2002a))
441 and D_{S} =0.75, we would estimate the coexisting inner core composition with about 1.7 wt% silicon
442 and 1.4 wt% sulfur (i.e., $c_{\mathrm{O}}^{\mathrm{sol}} \approx 0$, $c_{\mathrm{Si}}^{\mathrm{sol}} \approx c_{\mathrm{Si}}^{\mathrm{liq}} = 0.033$ and $c_{\mathrm{S}}^{\mathrm{sol}} = D_{\mathrm{S}} \times c_{\mathrm{S}}^{\mathrm{liq}} = 0.024$). Now with the
443 solid solution model provided by Li et al (2018), the obtained inner core density would agree well with
444 the PREM model by Dziewonski and Anderson (1981) with a relative error less than 0.5% and the
445 density jump would be around 0.65 g/cm^3. Note that we have not included hydrogen and carbon here
446 since the properties of their alloys with iron are much more unclear, although they may be potentially
447 important for explaining the seismic observations under certain circumstances (Li et al., 2018; Li et
448 al., 2019; Mashino et al., 2019; Umemoto and Hirose, 2020).

449 If the above core composition is plausibly in accord with geochemical and geophysical constraints,
450 then the roles of sulfur on the geodynamics and evolution of Earth's core should be minor. This may
451 be firstly inferred from its contribution to the density jump (0.65 g/cm^3 as mentioned above), which is
452 important for evaluating the gravitational energy to drive the geodynamo (Stacey and Stacey, 1999;
453 Gubbins et al., 2004). By "turning on" the partitioning coefficients stepwise (i.e., we inspect the
454 differences by switching the D_{O} and D_{S} from 1.0 to the expected values), the exclusion of oxygen in
455 the inner core and partitioning of sulfur account for about 0.37 g/cm^3 and 0.04 g/cm^3 respectively, in
456 addition to the contribution from freezing of iron (0.24 g/cm^3). It seems that the 6% contribution here
457 from sulfur cannot be entirely neglected, but this is highly likely to be an upper bound, since sulfur is
458 recently found to be less siderophile in the Earth's core (Suer et al., 2017).

459 Finally, the almost invariant partition coefficient from 250 GPa to higher pressures is interesting.
460 It would imply a constant partitioning of sulfur since the advent of the solid inner core to the time when

its radius expands to at least 1000 km larger as compared with the present size. This would support a stable stratification of sulfur in the Earth's inner core, since more sulfur precipitates in the inner core as the outer core sulfur concentration increases over time (Cottaar and Buffett, 2012; Deguen and Cardin, 2011). By contrast, without counting the *T*-*P* dependence of the relevant parameters in the equations of chemical potentials (as listed in Table 4 and analyzed in the previous section), recent numerical simulations propose continuously decreasing partition coefficient with time and find its destabilizing buoyancy effects on the stratification of the Earth's inner core (Gubbins et al., 2013; Labrosse, 2014; Lythgoe et al., 2015). Our results in this study show that these simulations may need to be re-evaluated at least for sulfur based on our results.

6. Concluding remarks

In this study, we derive new generation nonparametric interaction potentials for Fe-S systems applicable under Earth's core conditions. Based on machine learning techniques, these Gaussian Approximation Potentials are shown to reproduce the first principles simulation results with unprecedented accuracies, including the interatomic forces, local structures and, most importantly, the free energies that fundamentally govern all thermodynamic properties. With a similar approach, we will be able to derive accurate potentials for more complex systems (e.g., multicomponent systems including elements of Ni, O, S, Si, C, H), which are very difficult to investigate solely with first principles techniques due to the increase in the size of the phase space.

The substantial initial efforts of training the machine learning potentials provide a return in the remarkable efficiency in sampling the phase spaces of iron and its alloys under various temperature and pressure conditions. It is then possible for us to simulate free energies and predict phase behaviors

483 with fundamentally rigorous thermodynamic integration method within affordable computational cost.
484 In fact, to thoroughly sample the phase space around the liquidus and solidus of Fe-S solutions, we
485 have carried out over 500 independent atomistic simulations in this study, each with 180 atoms and at
486 least 20000 steps. The current implementation of the GAP models takes about 12 CPU seconds for
487 each step in each run. With parallel acceleration of 24 CPU cores for each run, we have managed to
488 accomplish the simulations within two months. As a comparison, the direct high precision DFT
489 simulations are about three orders of magnitudes slower than the GAP simulations.

490 As a first application of the framework mentioned above, we focus in this paper on the partition
491 coefficients of sulfur between the solid and liquid iron under Earth's core conditions. While the results
492 at ICB are in good agreements with early DFT simulations, we obtained the melting and partitioning
493 behaviors over the entire relevant *T-P* regime of the Earth's core. In particular, the invariance of
494 partition coefficients from 250 GPa to higher pressures found in this study provides new constraint on
495 the compositions and dynamics of Earth's inner and outer core.

496 Finally, since the phase behaviors of iron alloys are comprehensively complex, it should be noted
497 that much more endeavors are needed beyond our current efforts of predicting sulfur partitioning in
498 Fe-S binary systems under extreme conditions. The interplays of different impurities, the possible
499 stabilization of face-cubic-centered (*fcc*) or even body-cubic-centered (*bcc*) structures, the
500 immiscibility of liquid iron-alloying systems, the heterogeneities of the Earth's inner core, etc., can all
501 be important to estimate the partitioning of each specific light element across ICB in real Earth's core.
502 Interests in these issues would imply quickly growing demands of computations over broader phase
503 spaces. The high accuracy and efficiency gained by the framework proposed in this study would benefit
504 providing new constraints over all these issues.

505

506 **Supplementary materials**

507 Supplementary materials associated with this article can be found in: Appendix A that includes
508 several supporting figures for the main text; Appendix B that has the details for a GAP model training
509 with the QUIP code; Appendix C that includes the complementary research data and goes to the
510 Electronic Annex.

511

512 **Research data for this article**

513 Most data used in this article are given in Tables 2-4, complementary research data are included
514 in the Electronic Annex.

515

516 **ACKNOWLEDGMENTS** We thank helpful discussions with Drs. Tao Sun, Zuan Chen, Noam
517 Bernstein and Albert Bartok. We thank the assistance of surveying experimental data by Mr. Zhiming
518 Zhang. We are grateful to three anonymous reviewers for their constructive comments and to Marc
519 Blanchard and Jeffrey Catalano for their handling our manuscript. This work was supported by the
520 Strategic Priority Research Program (B) of Chinese Academy of Sciences (#XDB18000000), State
521 Key Development Program of Basic Research of China (2014CB845905) and the opening fund of
522 State Key Laboratory of Lunar and Planetary Sciences. Simulations were carried out on the
523 computational facilities in the Computer Simulation Lab of IGGCAS and Tianhe-2 at the National
524 Supercomputer Center of China (NSCC) in Guangzhou.

525

526

527 **References**

528 Alboussiere, T., Deguen, R., Melzani, M., 2010. Melting-induced stratification above the Earth's inner
529 core due to convective translation. Nature 466, 744-U749.
530 Alfè, D., 2009. PHON: A program to calculate phonons using the small displacement method.
531 Computer Physics Communications 180, 2622-2633.
532 Alfe, D., Gillan, M.J., 1998. First-principles simulations of liquid Fe-S under Earth's core conditions.
533 Phys Rev B 58, 8248-8256.
534 Alfe, D., Gillan, M.J., Price, G.D., 2002a. Ab initio chemical potentials of solid and liquid solutions
535 and the chemistry of the Earth's core. J Chem Phys 116, 7127-7136.
536 Alfe, D., Gillan, M.J., Price, G.D., 2002b. Complementary approaches to the ab initio calculation of
537 melting properties. J Chem Phys 116, 6170-6177.
538 Alfe, D., Price, G.D., Gillan, M.J., 2002c. Iron under Earth's core conditions: Liquid-state
539 thermodynamics and high-pressure melting curve from ab initio calculations. Phys Rev B 65.
540 Alfe, D., Gillan, M.J., Price, G.D., 2003. Thermodynamics from first principles: temperature and
541 composition of the Earth's core. Mineral Mag 67, 113-123.
542 Alfe, D., Gillan, M.J., Price, G.D., 2007. Temperature and composition of the Earth's core.
543 Contemporary Physics 48, 63-80.
544 Allegre, C.J., Poirier, J.P., Humler, E., Hofmann, A.W., 1995. The Chemical-Composition of the Earth.
545 Earth Planet Sc Lett 134, 515-526.
546 Anzellini, S., Dewaele, A., Mezouar, M., Loubeyre, P., Morard, G., 2013. Melting of Iron at Earth's
547 Inner Core Boundary Based on Fast X-ray Diffraction. Science 340, 464-466.
548 Aquilanti, G., Trapananti, A., Karandikar, A., Kantor, I., Marini, C., Mathon, O., Pascarelli, S., Boehler,

549 R., 2015. Melting of iron determined by X-ray absorption spectroscopy to 100 GPa. P Natl
550 Acad Sci USA 112, 12042-12045.
551 Arfken, G., 1985. Stirling's Series, in: Orlando, F. (Ed.), Mathematical Methods for Physicists, 3rd ed.
552 Academic Press, pp. 555-559.
553 Badro, J., Côté, A.S., Brodholt, J.P., 2014. A seismologically consistent compositional model of Earth's
554 core. Proceedings of the National Academy of Sciences.
555 Bartók, A.P., Kermode, J., Bernstein, N., Csányi, G., 2018. Machine Learning a General-Purpose
556 Interatomic Potential for Silicon. Phys Rev X 8, 041048.
557 Bartok, A.P., Gillan, M.J., Manby, F.R., Csanyi, G., 2013a. Machine-learning approach for one- and
558 two-body corrections to density functional theory: Applications to molecular and condensed
559 water. Phys Rev B 88.
560 Bartók, A.P., Kondor, R., Csányi, G., 2013b. On representing chemical environments. Phys Rev B 87,
561 184115.
562 Bartok, A.P., Payne, M.C., Kondor, R., Csanyi, G., 2010. Gaussian Approximation Potentials: The
563 Accuracy of Quantum Mechanics, without the Electrons. Phys Rev Lett 104.
564 Behler, J., Parrinello, M., 2007. Generalized neural-network representation of high-dimensional
565 potential-energy surfaces. Phys Rev Lett 98.
566 Belonoshko, A.B., Ahuja, R., Johansson, B., 2000. Quasi - Ab initio molecular dynamic study of Fe
567 melting. Phys Rev Lett 84, 3638-3641.
568 Ben-Naim, A., 2006. Molecular theory of solutions. Oxford University Press, New York.
569 Birch, F., 1964. Density and composition of mantle and core. Journal of Geophysical Research 69,
570 4377-&.

571 Boehler, R., 1993. Temperatures in the Earth's core from melting-point measurements of iron at high
572 static pressures. Nature 363, 534-536.
573 Brown, J.M., McQueen, R.G., 1986. Phase transition, Grüneisen parameter, and elasticity for shocked
574 iron between 77 GPa and 400 GPa. J Geophys Res 91.
575 Cao, A., Romanowicz, B., 2004. Constraints on density and shear velocity contrasts at the inner core
576 boundary. Geophys J Int 157, 1146-1151.
577 Ceriotti, M., Willatt, M.J., Csányi, G., 2018. Machine Learning of Atomic-Scale Properties Based on
578 Physical Principles, in: Andreoni, W., Yip, S. (Eds.), Handbook of Materials Modeling :
579 Methods: Theory and Modeling. Springer International Publishing, Cham, pp. 1-27.
580 Chen, B., Li, J., Hauck, S.A., 2008. Non-ideal liquidus curve in the Fe-S system and Mercury's
581 snowing core. Geophys Res Lett 35.
582 Chudinovskikh, L., Boehler, R., 2007. Eutectic melting in the system Fe–S to 44 GPa. Earth Planet Sc
583 Lett 257, 97-103.
584 Cote, A.S., Vocadlo, L., Brodholt, J.P., 2008. Light elements in the core: Effects of impurities on the
585 phase diagram of iron. Geophys Res Lett 35, L05306.
586 Cottaar, S., Buffett, B., 2012. Convection in the Earth's inner core. Phys Earth Planet In 198, 67-78.
587 Deguen, R., Cardin, P., 2011. Thermochemical convection in Earth's inner core. Geophys J Int 187,
588 1101-1118.
589 Deringer, V.L., Caro, M.A., Csanyi, G., 2019. Machine Learning Interatomic Potentials as Emerging
590 Tools for Materials Science. Adv Mater.
591 Deringer, V.L., Csanyi, G., Proserpio, D.M., 2017. Extracting Crystal Chemistry from Amorphous
592 Carbon Structures. ChemPhysChem 18, 873-877.

593 Dorner, F., Sukurma, Z., Dellago, C., Kresse, G., 2018. Melting Si: Beyond Density Functional Theory.
594 Phys Rev Lett 121.
595 Dragoni, D., Daff, T.D., Csanyi, G., Marzari, N., 2018. Achieving DFT accuracy with a machine-
596 learning interatomic potential: Thermomechanics and defects in bcc ferromagnetic iron. Phys.
597 Rev. Mater. 2.
598 Drautz, R., 2019. Atomic cluster expansion for accurate and transferable interatomic potentials. Phys
599 Rev B 99.
600 Dziewonski, A.M., Anderson, D.L., 1981. Preliminary reference Earth model. Phys Earth Planet Inter
601 25.
602 Gavryushkin, P.N., Popov, Z.I., Litasov, K.D., Belonoshko, A.B., Gavryushkin, A., 2016. Stability of
603 B2-type FeS at Earth's inner core pressures. Geophys Res Lett 43, 8435-8440.
604 Gubbins, D., Alfe, D., Masters, G., Price, G.D., Gillan, M., 2004. Gross thermodynamics of two-
605 component core convection. Geophys J Int 157, 1407-1414.
606 Gubbins, D., Alfe, D., Davies, C.J., 2013. Compositional instability of Earth's solid inner core.
607 Geophys Res Lett 40, 1084-1088.
608 Hirose, K., Labrosse, S., Hernlund, J., 2013. Composition and State of the Core. Annual Review of
609 Earth and Planetary Sciences, Vol 41 41, 657-691.
610 Kamada, S., Ohtani, E., Terasaki, H., Sakai, T., Miyahara, M., Ohishi, Y., Hirao, N., 2012. Melting
611 relationships in the Fe-Fe3S system up to the outer core conditions. Earth Planet Sc Lett 359,
612 26-33.
613 Kamada, S., Terasaki, H., Ohtani, E., Sakai, T., Kikegawa, T., Ohishi, Y., Hirao, N., Sata, N., Kondo,
614 T., 2010. Phase relationships of the Fe-FeS system in conditions up to the Earth's outer core.

615 Earth Planet Sc Lett 294, 94-100.
616 Komabayashi, T. and Fei, Y.W., 2010. Internally consistent thermodynamic database for iron to the
617 Earth's core conditions. J Geophys Res-Sol Ea 115, B03202.
618 Kresse, G., Joubert, D., 1999. From ultrasoft pseudopotentials to the projector augmented-wave
619 method. Phys Rev B 59, 1758-1775.
620 Kung, S.Y., 2014. Kernel Methods and Machine Learning. Cambridge University Press, Cambridge.
621 Labrosse, S., 2014. Thermal and compositional stratification of the inner core. Comptes Rendus
622 Geoscience 346, 119-129.
623 Laio, A., Bernard, S., Chiarotti, G.L., Scandolo, S., Tosatti, E., 2000. Physics of iron at Earth's core
624 conditions. Science 287.
625 Li, J., Fei, Y., 2007. 2.14 - Experimental Constraints on Core Composition, in: Holland, H.D., Turekian,
626 K.K. (Eds.), Treatise on Geochemistry. Pergamon, Oxford, pp. 1-31.
627 Li, J., Fei, Y., Mao, H.K., Hirose, K., Shieh, S.R., 2001. Sulfur in the Earth's inner core. Earth Planet
628 Sc Lett 193, 509-514.
629 Li, Y., Vočadlo, L., Brodholt, J.P., 2018. The elastic properties of hcp-Fe alloys under the conditions
630 of the Earth's inner core. Earth Planet Sc Lett 493, 118-127.
631 Li, Y., Vocadlo, L., Alfe, D., Brodholt, J., 2019. Carbon Partitioning Between the Earth's Inner and
632 Outer Core. J Geophys Res-Sol Ea 124, 12812-12824.
633 Lythgoe, K.H., Rudge, J.F., Neufeld, J.A., Deuss, A., 2015. The feasibility of thermal and
634 compositional convection in Earth's inner core. Geophys J Int 201, 764-782.
635 Ma, Z.T., 2001. Thermodynamic description for concentrated metallic solutions using interaction
636 parameters. Metall Mater Trans B 32, 87-103.

637 Mashino, I., Miozzi, F., Hirose, K., Morard, G., Sinmyo, R., 2019. Melting experiments on the Fe-C
638 binary system up to 255 GPa: Constraints on the carbon content in the Earth's core. Earth Planet
639 Sc Lett 515, 135-144.
640 Masters, G., Gubbins, D., 2003. On the resolution of density within the Earth. Phys Earth Planet In
641 140, 159-167.
642 McDonough, W.F., 2003. Compositional Model for the Earth's Core, in: Holland, H.D., Turekian, K.K.
643 (Eds.), Treatise on Geochemistry. Pergamon, Oxford, pp. 547-568.
644 McDonough, W.F., Sun, S.S., 1995. The Composition of the Earth. Chem Geol 120, 223-253.
645 Mermin, N.D., 1965. Thermal Properties of Inhomogeneous Electron Gas. Phys Rev 137, 1441-&.
646 Morard, G., Andrault, D., Antonangeli, D., Bouchet, J., 2014. Properties of iron alloys under the Earth's
647 core conditions. Comptes Rendus Geoscience 346, 130-139.
648 Morard, G., Andrault, D., Guignot, N., Sanloup, C., Mezouar, M., Petitgirard, S., Fiquet, G., 2008. In
649 situ determination of Fe-Fe3S phase diagram and liquid structural properties up to 65 GPa.
650 Earth Planet Sc Lett 272, 620-626.
651 Morard, G., Boccato, S., Rosa, A.D., Anzellini, S., Miozzi, F., Henry, L., Garbarino, G., Mezouar, M.,
652 Harmand, M., Guyot, F., Boulard, E., Kantor, I., Irifune, T., Torchio, R., 2018. Solving
653 Controversies on the Iron Phase Diagram Under High Pressure. Geophys Res Lett 45, 11074-
654 11082.
655 Mori, Y., Ozawa, H., Hirose, K., Sinmyo, R., Tateno, S., Morard, G., Ohishi, Y., 2017. Melting
656 experiments on Fe-Fe3S system to 254 GPa. Earth Planet Sc Lett 464, 135-141.
657 Nguyen, J.H., Holmes, N.C., 2004. Melting of iron at the physical conditions of the Earth's core. Nature
658 427, 339-342.

659 Perdew, J.P., Burke, K., Ernzerhof, M., 1996. Generalized gradient approximation made simple. Phys
660 Rev Lett 77, 3865-3868.
661 Poirier, J.-P., 1994. Light elements in the Earth's outer core: A critical review. Phys Earth Planet In 85,
662 319-337.
663 Rupp, M., 2015. Special issue on machine learning and quantum mechanics. Int J Quantum Chem 115,
664 1003-1004.
665 Shapeev, A., 2017. Accurate representation of formation energies of crystalline alloys with many
666 components. Comp Mater Sci 139, 26-30.
667 Sinmyo, R., Hirose, K. and Ohishi, Y., 2019. Melting curve of iron to 290 GPa determined in a
668 resistance-heated diamond-anvil cell. Earth Planet Sc Lett 510, 45-52.
669 Stacey, F.D., Stacey, C.H.B., 1999. Gravitational energy of core evolution: implications for thermal
670 history and geodynamo power. Phys Earth Planet In 110, 83-93.
671 Stewart, A.J., Schmidt, M.W., van Westrenen, W., Liebske, C., 2007. Mars: A new core-crystallization
672 regime. Science 316, 1323-1325.
673 Suer, T.A., Siebert, J., Remusat, L., Menguy, N., Fiquet, G., 2017. A sulfur-poor terrestrial core inferred
674 from metal-silicate partitioning experiments. Earth Planet Sc Lett 469, 84-97.
675 Sun, T., Brodholt, J.P., Li, Y.G., Vocadlo, L., 2018. Melting properties from ab initio free energy
676 calculations: Iron at the Earth's inner-core boundary. Phys Rev B 98.
677 Szlachta, W.J., Bartok, A.P., Csanyi, G., 2014. Accuracy and transferability of Gaussian approximation
678 potential models for tungsten. Phys Rev B 90.
679 Terasaki, H., Kamada, S., Sakai, T., Ohtani, E., Hirao, N., Ohishi, Y., 2011. Liquidus and solidus
680 temperatures of a Fe-O-S alloy up to the pressures of the outer core: Implication for the thermal

681 structure of the Earth's core. Earth Planet Sc Lett 304, 559-564.
682 Umemoto, K., Hirose, K., 2020. Chemical compositions of the outer core examined by first principles
683 calculations. Earth Planet Sc Lett 531, 8, 116009.
684 Veit, M., Jain, S.K., Bonakala, S., Rudra, I., Hohl, D., Csanyi, G., 2019. Equation of State of Fluid
685 Methane from First Principles with Machine Learning Potentials. Journal of Chemical Theory
686 and Computation 15, 2574-2586.
687 White, v.W.M., 2013. Geochemistry. Wiley-Blackwell.
688 Widom, M., Huhn, W.P., Maiti, S., Steurer, W., 2014. Hybrid Monte Carlo/Molecular Dynamics
689 Simulation of a Refractory Metal High Entropy Alloy. Metall Mater Trans A 45A, 196-200.
690 Wood, B.J., Walter, M.J., Wade, J., 2006. Accretion of the Earth and segregation of its core. Nature
691 441, 825-833.
692 Yokoo, S., Hirose, K., Sinmyo, R., Tagawa, S., 2019. Melting Experiments on Liquidus Phase
693 Relations in the Fe-S-O Ternary system Under Core Pressures. Geophys Res Lett 46, 5137-
694 5145.
695 Yoo, C.S., Holmes, N.C., Ross, M., Webb, D.J., Pike, C., 1993. Shock temperatures and melting of
696 iron at earth core conditions. Phys Rev Lett 70, 3931-3934.
697

698

699 Table 1. Hyper-parameters for training the GAP models

GAP software version		1527075646
Atomic environment kernel		SOAP
r_{cut}		5.0 Å
r_Δ		1.0 Å
σ^{energy} default		0.002 eV/atom
σ^{force} default		0.05 eV/Å
σ^{virial} default		0.05 eV/atom
σ_w		1.0 eV
σ_{atom}		0.5 Å
ξ		2
n_{max}		8
l_{max}		8
Representative environments		4000
Sparse method		CUR
e_0 (eV)	4000 K	Fe:-2.70299745:S:-1.33262117
	5000 K	Fe:-3.36146432:S:-1.66173158
	6000 K	Fe:-4.02404024:S:-1.99085328
	7000 K	Fe:-4.69558982:S:-2.32005321

700

701

702

703 Table 2. Simulated Helmoholtz free energies (F) of pure iron

T (K)	P (GPa)	Liquid		Solid	
		V (cm^3/mol)	F (kJ/mol)	V (cm^3/mol)	F (kJ/mol)
4000	100	5.4277	-1040.49±1.16	5.2783	-1024.35±0.46
5000	190	4.8171	-1061.59±1.25	4.7146	-1041.26±0.48
6000	330	4.2473	-1024.00±1.27	4.1721	-1001.65±0.59
7000	400	4.0714	-1069.38±1.29	4.0072	-1041.69±0.94

704

705

706

707 Table 3. Simulated Helmoholtz free energies (F) of Fe-S alloys

T (K)	Phase	c_S	V (cm^3/mol)	P (GPa)	F (kJ/mol)
4000	Liquid	0.05	5.4239	100.07±0.15	-1028.15±0.81
		0.10	5.4245	100.12±0.14	-1012.87±0.77
		0.15	5.4330	99.64±0.30	-996.13±1.11
		0.20	5.4462	99.80±0.24	-978.22±0.84
4000	Solid	0.05	4.3603	249.61±0.05	-858.26±0.42
		0.10	4.3896	241.92±0.06	-847.21±0.42
		0.15	4.3819	243.42±0.11	-826.18±0.50
		0.20	4.3893	243.87±0.16	-804.96±0.47
5000	Liquid	0.05	4.8086	189.94±0.20	-1047.60±0.92
		0.10	4.7990	190.30±0.17	-1030.67±1.04
		0.15	4.7968	190.14±0.14	-1012.67±1.10
		0.20	4.7955	189.88±0.20	-992.56±0.95
5000	Solid	0.05	4.1225	329.36±0.05	-878.78±0.30
		0.10	4.1131	330.06±0.07	-859.37±0.31
		0.15	4.1139	328.59±0.12	-840.38±0.49
		0.20	4.1153	327.80±0.09	-820.03±0.42
6000	Liquid	0.05	4.2395	329.51±0.20	-1010.93±1.67
		0.10	4.2268	330.44±0.37	-992.56±1.03
		0.15	4.2195	330.32±0.17	-973.51±1.21
		0.20	4.2134	329.73±0.19	-951.78±1.26
6000	Solid	0.05	4.1636	329.27±0.10	-987.78±0.54
		0.10	4.1807	322.99±0.10	-977.19±0.43
		0.15	4.1692	325.74±0.20	-954.92±0.47

708

709 Table 4. Parameters for chemical potential and free energy at various T-P conditions (Eqns. (8-10))

T	P	Liquid			Solid		
(K)	(GPa)	μ^0_{Fe} (kJ/mol)	$\mu^{\ddagger}_S$ (kJ/mol)	λ_S (kJ/mol)	μ^0_{Fe} (kJ/mol)	$\mu^{\ddagger}_S$ (kJ/mol)	λ_S (kJ/mol)
4000	150	-237.94±1.16	116.43±1.11	463.67±12.04	-243.08±0.46	182.49±7.23	271.38±92.00
4000	200	3.66±1.16	349.75±1.97	496.38±21.30	-6.50±0.46	415.01±5.81	317.29±73.86
4000	250	231.91±1.16	566.66±2.19	525.74±23.73	217.31±0.46	629.14±7.19	405.41±91.53
5000	250	132.88±1.25	518.53±6.73	322.69±72.17	128.11±0.48	542.03±2.73	365.32±31.77
5000	300	353.11±1.25	726.27±6.97	361.62±74.72	344.14±0.48	746.94±2.80	442.20±32.65
5000	330	480.79±1.25	844.73±6.68	393.97±71.60	469.52±0.48	865.41±2.84	478.25±33.07
6000	330	377.61±1.27	766.37±5.51	375.25±61.15	375.48±0.50	779.01±6.12	510.14±88.08

710

711 **Figure Captions**

712 Figure 1. Comparisons of atomic forces (a), stresses (b) and energies (c) between DFT and the GAP
713 model at 6000 K. 1178 energies, 3534 normal stress components and 344928 forces are
714 included in these comparisons. The red dashed lines are guides for perfect matches. In the
715 inset plots, we show the cumulative probability distribution of force component errors and
716 energy errors (relative to reference DFT calculations).

717 Figure 2. Comparisons of the radial distribution functions of a Fe-S liquid (6000 K, 237 GPa and
718 c_S=0.21) from the DFT simulation (full lines) and that with the GAP models (broken lines).

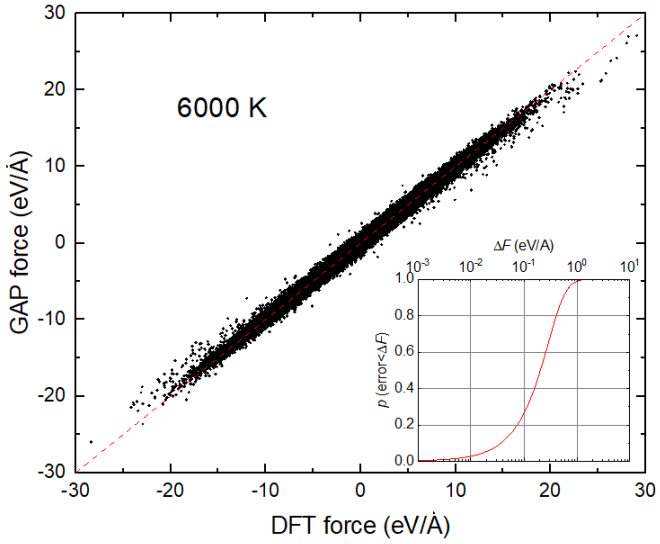
719 Figure 3. Equation of state (a) and free energy profiles (b) of pure iron at 5000 K. In (a), the filled
720 symbols are simulation results with GAP models without phase transition over the whole
721 trajectories, while the open symbols are the statistical averages after the phase changes. The
722 dashed lines are Birch-Murnaghan equations regressed from the filled symbols. The solid
723 lines are those provided by Komabayashi and Fei (2010) based on experimental data. In the
724 upper plot of (b), the line is the Gibbs free energy change ($\Delta G=G_{solid}-G_{liquid}$) and light-blue
725 shadow marks the uncertainty.

726 Figure 4. Melting of iron from different approaches. Symbols and lines in blue are the experimental
727 data, whole those in black are the previous simulation results. The red squares are the
728 melting points from this study with GAP models.

729 Figure 5. Gibbs free energies of Fe-S solutions under 4000 K, 250 GPa (a) and 6000 K, 330 GPa (b).
730 The filled black and red symbols are the free energies for liquids and solids, respectively.
731 The lines are regressions with Eqn. (10) based on these data (relevant parameters are listed
732 in Table 4). Open blue circles are the results for solid solutions through simple molecular

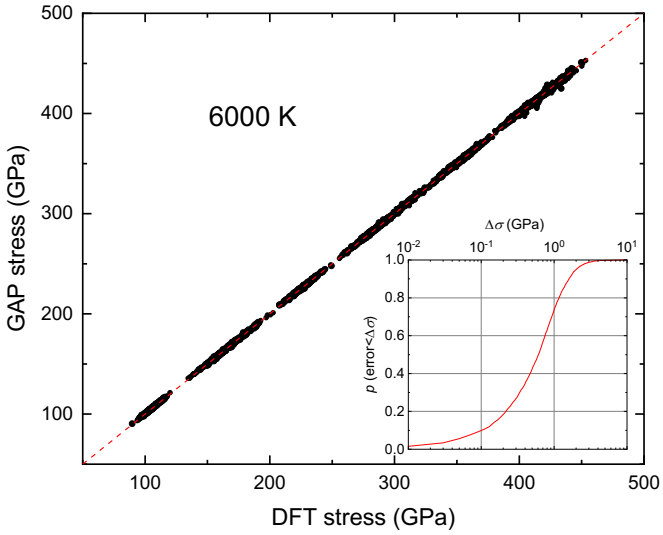
733 dynamics samplings, as compared with the filled red circles through hybrid Monte
734 Carlo/molecular dynamics samplings (see the text). The uncertainties of the free energies
735 are generally smaller than the sizes of the symbols.
736 Figure 6. Effects of Monte Carlo swapping on the samplings of thermodynamic properties and
737 structures of the Fe-S solid solution. In (a), we show the propagations of total energy (U),
738 pressure (P) and number of nearest-neighbored S-S pairs ($NN_{S\text{-}S}$) during the trajectories of
739 the two simulations (one with simple MD, in green; the other with hybrid Monte Carlo and
740 MD, or MD+MC, in black) at 4000 K, 4.3893 cm^3/mol, c_S=0.20. The broken lines are
741 equilibrium averaged values (i.e., discarding the first 5000 fs as pre-equilibrium stage) from
742 both simulations (MD, in blue; MD+MC, in red) for better guides of the fluctuations. In (b),
743 we compare the radial distribution functions of different pairs from the two simulations.
744 Figure 7. Partition coefficients of sulfur between solid and liquid iron. Filled symbols are those from
745 this study at different temperatures and pressures. The additional red crossed circle at 6000
746 K and 330 GPa is a benchmark for simple MD simulations with random solid solutions.
747 Open symbols are data from previous experiments (Kamada et al., 2012; Kamada et al.,
748 2010; Li et al., 2001; Mori et al., 2017; Stewart et al., 2007; Yokoo et al., 2019) and
749 simulations (Alfe et al., 2002a). Note that all the experimental data have been carefully
750 converted to the ratio of mole faction to be consistent with simulation results.
751
752

753



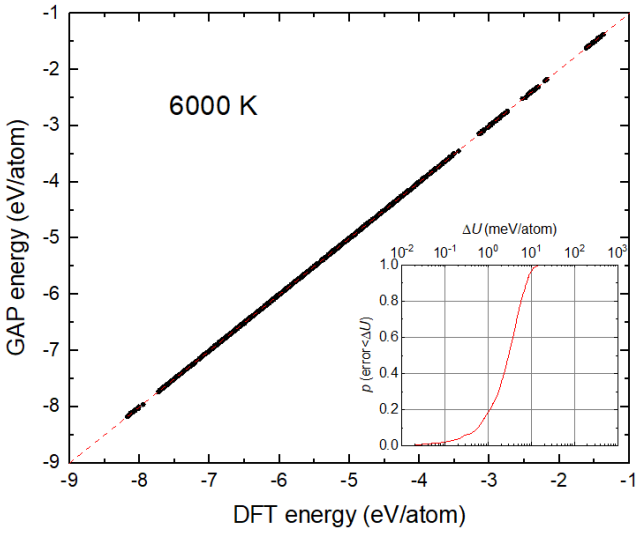
754

755 (a)



756

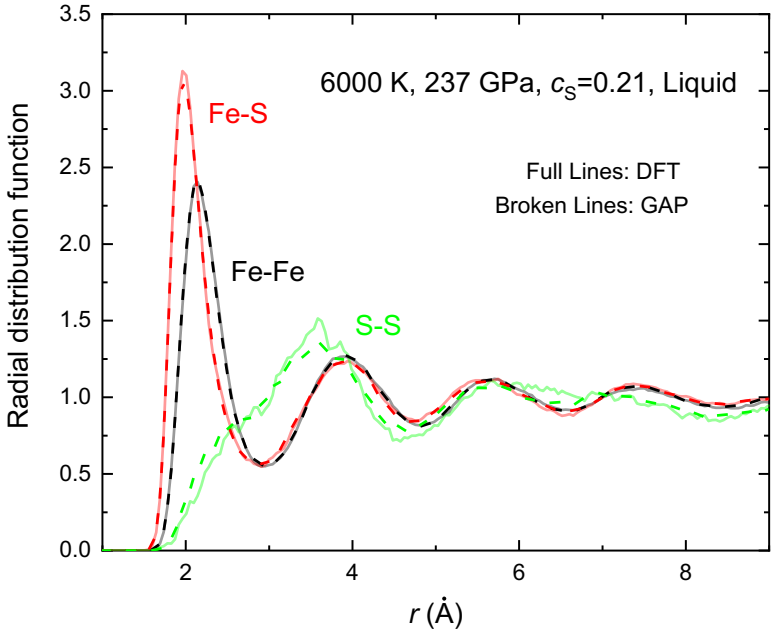
757 (b)


6000 K
GAP energy (eV/atom)
DFT energy (eV/atom)
ΔU (meV/atom)
p (error<ΔU)

(c)

Figure 1

763

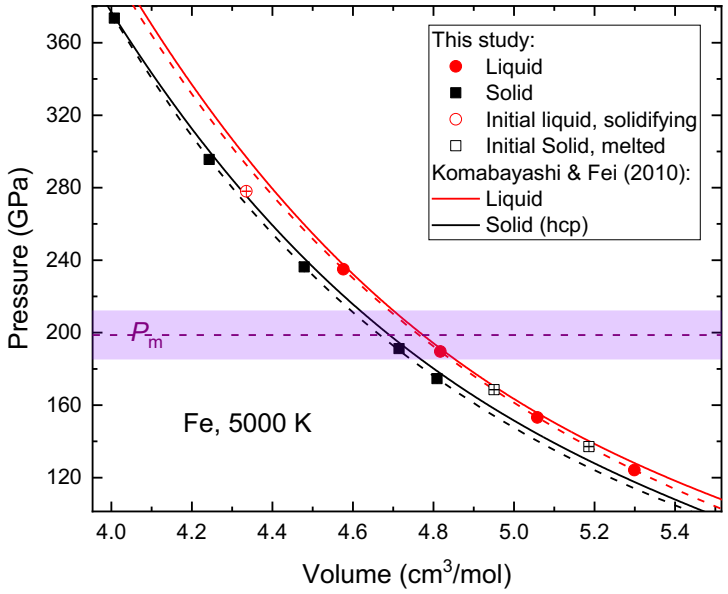


764

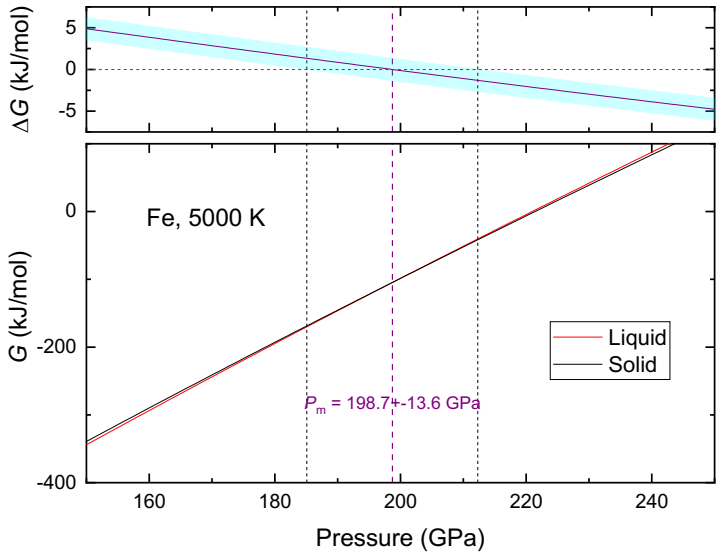
765 Figure 2

766

767


This study:
Liquid
Solid
Initial liquid, solidifying
Initial Solid, melted
Komabayashi & Fei (2010):
Liquid
Solid (hcp)
P_m
Fe, 5000 K
Pressure (GPa)
Volume (cm^3/mol)

(a)

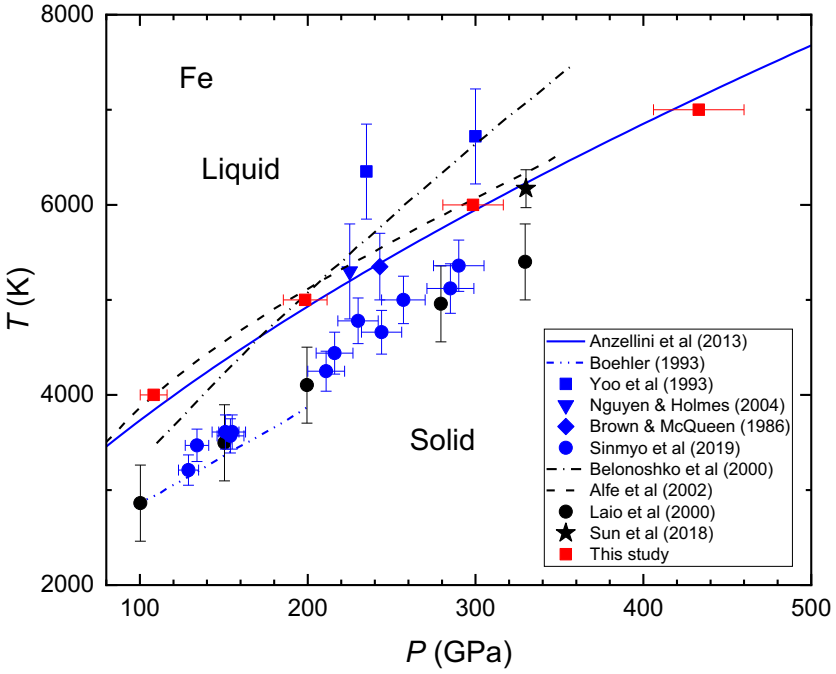

ΔG (kJ/mol)
G (kJ/mol)
Fe, 5000 K
Liquid
Solid
P_m = 198.7+-13.6 GPa
Pressure (GPa)

(b)

Figure 3

775

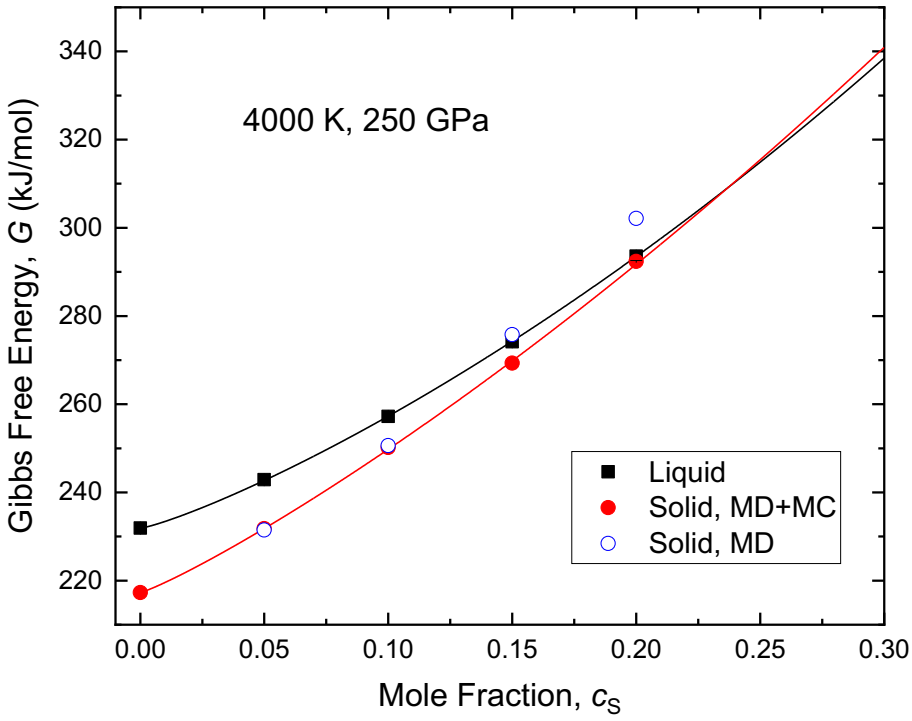
776



777

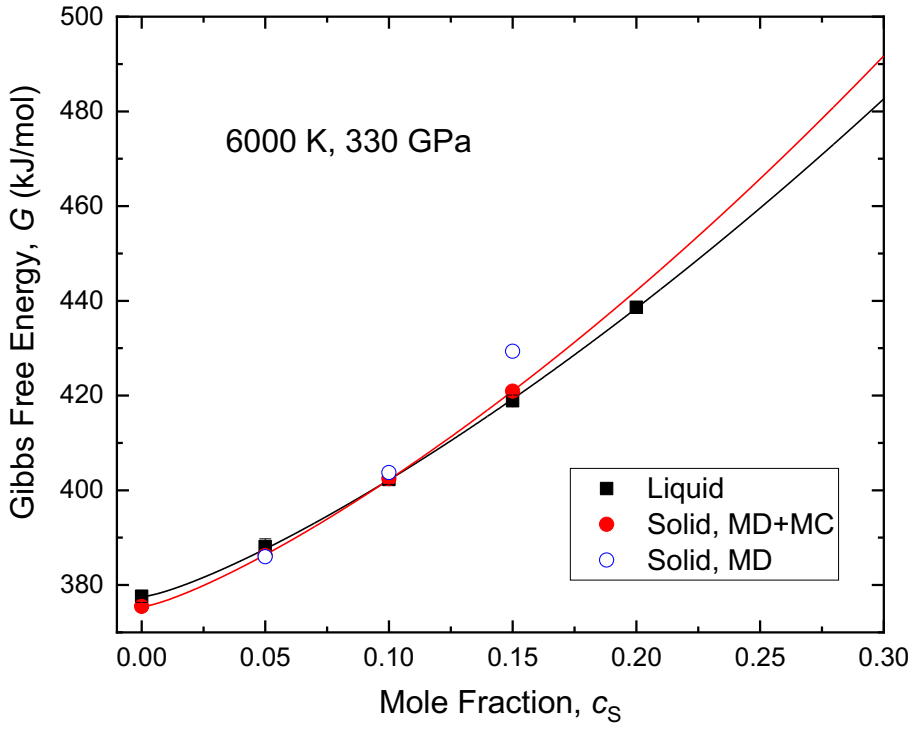
778 Figure 4

779



780

781 (a)

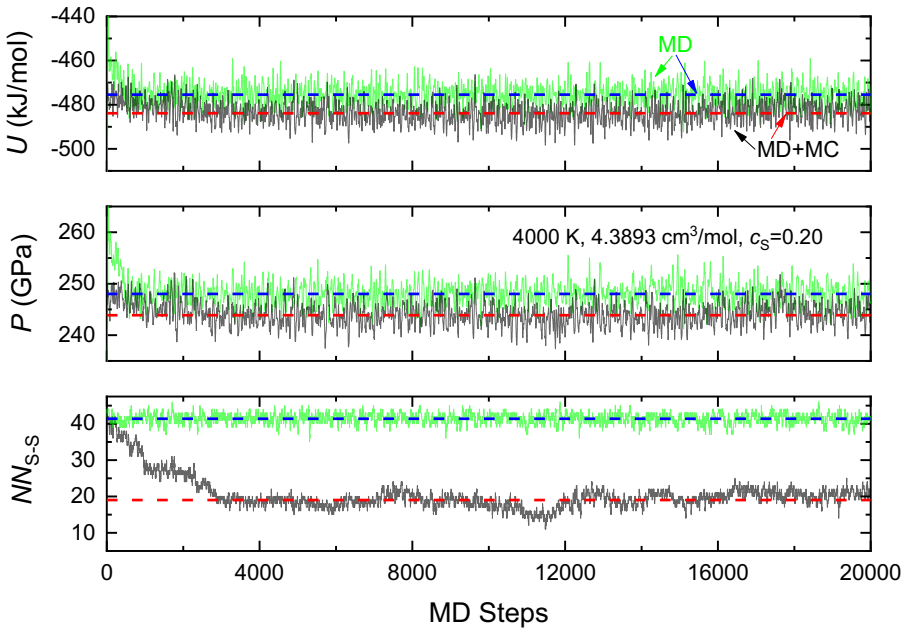


782

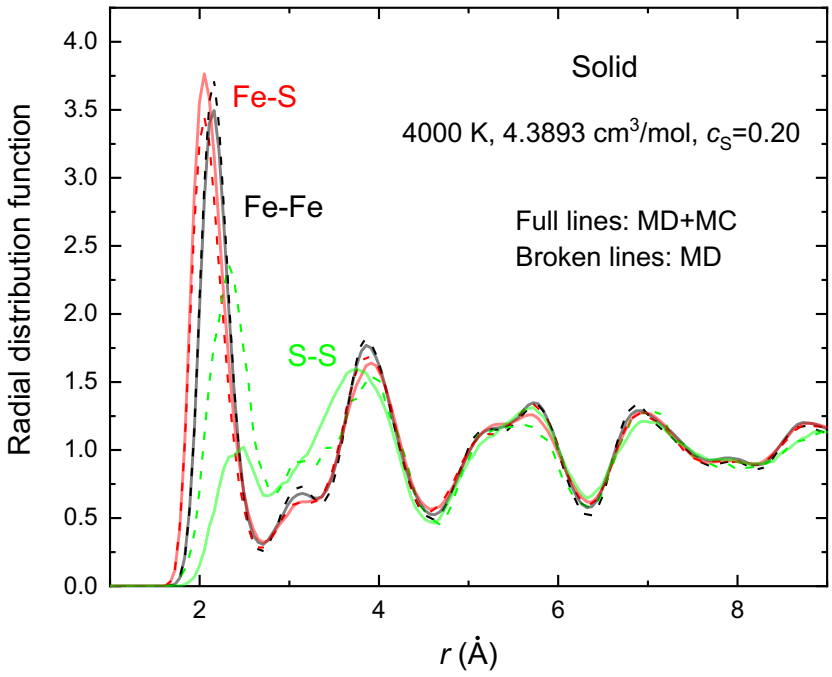
783 (b)

784 Figure 5

785

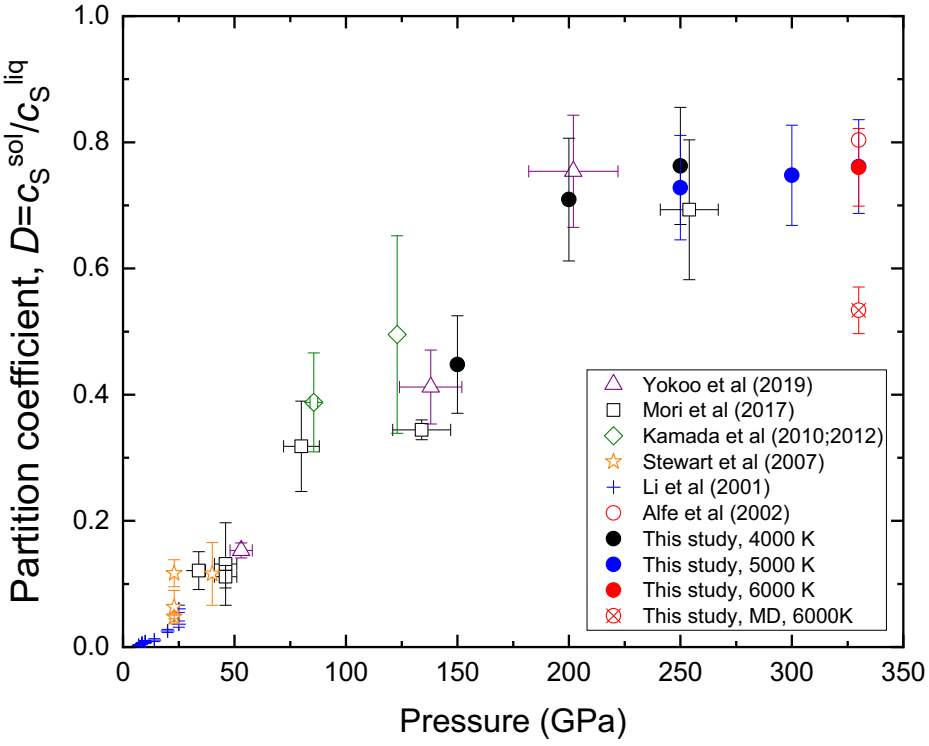

U (kJ/mol)
-440
-460
-480
-500
MD
MD+MC
P (GPa)
260
250
240
4000 K, 4.3893 cm3/mol, cS=0.20
NNS-S
40
30
20
10
0
4000
8000
12000
16000
20000
MD Steps

(a)


Radial distribution function
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Solid
Fe-S
4000 K, 4.3893 cm3/mol, cS=0.20
Fe-Fe
Full lines: MD+MC
Broken lines: MD
S-S
2
4
6
8
r (Å)

(b)

Figure 6



794

795 Figure 7

796

797