



# Partitioning of sulfur between solid and liquid iron under Earth's core conditions: Constraints from atomistic simulations with machine learning potentials

Zhigang Zhang<sup>a,b,c,\*</sup>, Gábor Csányi<sup>d</sup>, Dario Alfè<sup>e,f</sup>

<sup>a</sup> Key Laboratory of Earth and Planetary Physics, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China

<sup>b</sup> State Key Laboratory of Lunar and Planetary Sciences (Macau University of Science and Technology), Macau, China

<sup>c</sup> College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

<sup>d</sup> Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, United Kingdom

<sup>e</sup> Department of Earth Sciences and London Centre for Nanotechnology, University College London, Gower Street, London WC1E 6BT, United Kingdom

<sup>f</sup> Dipartimento di Fisica "Ettore Pancini," Università degli Studi di Napoli "Federico II," Complesso di Monte S. Angelo, via Cinthia, 80126 Napoli, Italy

Received 22 October 2019; accepted in revised form 19 March 2020; available online 4 April 2020

## 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  $\sim 250$  GPa, the partition coefficient is found to be around  $0.75 \pm 0.07$  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 of the Earth's core should be minor.

© 2020 Elsevier Ltd. All rights reserved.

**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 (Birch, 1964; Allegre et al., 1995; McDonough and Sun, 1995). The prime candidates for these light elements include sulfur, oxygen, silicon, carbon, and hydrogen (Poirier, 1994; Li and Fei, 2007; Hirose et al., 2013). It is recognized that they

\* Corresponding author at: Key Laboratory of Earth and Planetary Physics, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China.

E-mail address: [zgzhang@mail.iggcas.ac.cn](mailto:zgzhang@mail.iggcas.ac.cn) (Z. Zhang).

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 (Masters and Gubbins, 2003; Cao and Romanowicz, 2004), 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 (Li et al., 2001; Chudinovskikh and Boehler, 2007; Stewart et al., 2007; Chen et al., 2008; Morard et al., 2008; Kamada et al., 2010; Terasaki et al., 2011; Kamada et al., 2012; Morard et al., 2014; Mori et al., 2017; 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 (Alfè 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

$$E = \sum_i \varepsilon(\mathbf{q}_i) \quad (1)$$

where the short-ranged local atomic energy  $\varepsilon$  is assumed to depend explicitly on the positions of the atoms within a sphere of radius  $r_{\text{cut}}$  centered on atom  $i$  and  $\mathbf{q}_i$  is a vector representing its local atomic environment. In the GAP framework, Gaussian process regression is used to model  $\varepsilon$ ,

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

Given two atoms  $i$  and  $j$ , the kernel function  $K(\mathbf{q}_i, \mathbf{q}_j)$  is the expected covariance of the respective local atomic energies  $\varepsilon(\mathbf{q}_i)$  and  $\varepsilon(\mathbf{q}_j)$ , 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),

$$K(\mathbf{q}_i, \mathbf{q}_j) = \sigma_w^2 |\hat{\mathbf{q}}_i \cdot \hat{\mathbf{q}}_j|^\xi \quad (3)$$

where the descriptor  $\hat{\mathbf{q}}_i$  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  $\xi$  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:

$$\mu_S^{\text{liq}}(T, P, c_S^{\text{liq}}) = \mu_S^{\text{sol}}(T, P, c_S^{\text{sol}}) \quad (4)$$

$$\mu_{\text{Fe}}^{\text{liq}}(T, P, c_S^{\text{liq}}) = \mu_{\text{Fe}}^{\text{sol}}(T, P, c_S^{\text{sol}}) \quad (5)$$

where  $\mu_S$  and  $\mu_{\text{Fe}}$  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  $\mu_S$  diverges logarithmically in the low-concentration limit ( $c_S \rightarrow 0$ ), at each temperature and pressure it is useful to express  $\mu_S$  as (Alfe et al., 2002a)

$$\mu_S = k_B T \ln c_S + \bar{\mu}_S(c_S) \quad (6)$$

where  $\bar{\mu}_S(c_S)$  is well behaved for all concentrations. While many models have been proposed for  $\bar{\mu}_S(c_S)$  (or its equivalences), such as various symmetric/asymmetric regular solution models (White, 2013) or those proposed by Ma (2001),

the following expansion for  $\bar{\mu}_S(c_S)$  is simple but practically meaningful

$$\bar{\mu}_S(c_S) = \mu_S^\ddagger + \lambda_S c_S + O(c_S^2) \quad (7)$$

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

$$\mu_S = k_B T \ln c_S + \mu_S^\ddagger + \lambda_S c_S \quad (8)$$

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

$$\mu_{\text{Fe}} = \mu_{\text{Fe}}^0 + (k_B T + \lambda_S) \ln(1 - c_S) + \lambda_S c_S \quad (9)$$

where  $\mu_{\text{Fe}}^0$  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

$$\begin{aligned} G &= N_S \mu_S + N_{\text{Fe}} \mu_{\text{Fe}} = N [c_S \mu_S + (1 - c_S) \mu_{\text{Fe}}] \\ &= N \{ [c_S \ln c_S + (1 - c_S) \ln(1 - c_S)] k_B T + (1 - c_S) \mu_{\text{Fe}}^0 \\ &\quad + c_S \mu_S^\ddagger + [c_S + (1 - c_S) \ln(1 - c_S)] \lambda_S \} \end{aligned} \quad (10)$$

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

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

## 2.3. Free energy calculations

In this study, we calculated the Helmholtz free energies ( $F$ ) using thermodynamic integration, which rigorously

Table 1  
Hyper-parameters for training the GAP models.

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

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

### 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 Eq. (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 endpoints, 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 Eq. (11) becomes

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

$k$  in Eqs. (13) and (14) is selected to be 0.8, which avoids the endpoint singularity since the  $\overline{\Delta U}(\lambda(x_i)) \lambda(x_i)^k$  can be safely set to be zero when  $\lambda = 0$  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  $\omega_{x_i}$  in Eq. (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 Eq. (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 and quasi-harmonic approximation theory (Alfè, 2009). From the harmonic system to the target system, since the energetic changes are small and vary smoothly with the integration variable (as shown in Fig. A.1(b) of Appendix A), we use the simpler three-point Gauss-Legendre quadrature by the following linear transformation

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

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

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

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

## 3. SIMULATION DETAILS

### 3.1. First principles simulations

To train the GAP models, we carried out a number of density functional theory (DFT) simulations with VASP, with the projector-augmented-wave method (Kresse and Joubert, 1999). We use the PBE form of Generalized Gradient Approximation (GGA) (Perdew et al., 1996) with valence electrons of 16 (valence configuration  $3s^2 3p^6 3d^7 4s^1$ ) for iron, which has been demonstrated to closely resemble the all-electron potential and be important for obtaining accurate melting properties (Sun et al., 2018), and 6 ( $3s^2 3p^4$ ) for sulfur.

Extensive molecular dynamics (MD) simulations have been deployed to sufficiently sample the *T-P-X* space spanning over 0–25 at. % sulfur concentrations, from 1000 K to 7000 K and from about 50 GPa to 500 GPa for the solid phases, and from 4000 K to 10000 K and from about 100 GPa to 500 GPa for the liquid phases. Overall 243 simulations have been carried out for the solid phases and 215

simulations for the liquid phases. The solid solutions were initiated with hexagonal close-packed (*hcp*) structures with 96 atoms in the supercells ( $4 \times 4 \times 3$  extension of the unit-cell). At high enough temperatures, the solids would always be melted and the final configurations of these runs were adopted to initialize liquid simulations. MD trajectories were propagated in the *NVT* ensemble with the Nosé thermostat for  $3 \sim 6$  ps. At each time step (1 fs interval), the electronic structure is calculated at the Brillouin zone center with an energy cutoff (ENCUT) of 500 eV and iteration convergence criterion (EDIFF) of  $10^{-6}$  eV.

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

### 3.2. Training of GAP models

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

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

### 3.3. Atomistic simulations with the GAP models

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

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

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

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

Finally, lattice dynamics simulations were deployed with PHON through small displacement 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  $0.01 \text{ \AA}$  for each atom. A dense  $30 \times 30 \times 30$  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 344,928 data retrieved from both simulations at 6000 K,

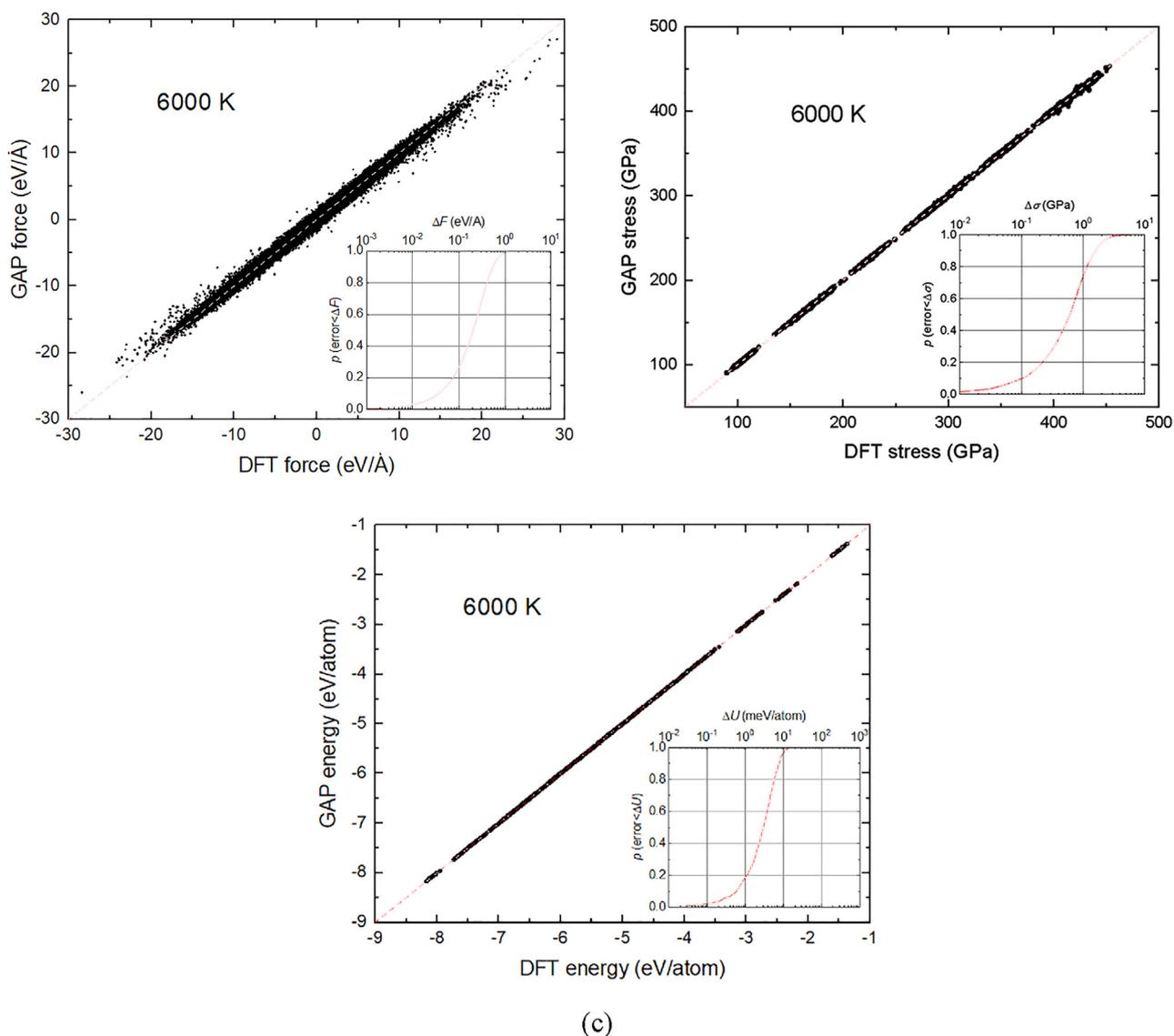


Fig. 1. Comparisons of atomic forces (a), stresses (b) and energies (c) between DFT and the GAP model at 6000 K. 1178 energies, 3534 normal stress components and 344,928 forces are included in these comparisons. The red dashed lines are guides for perfect matches. In the inset plots, we show the cumulative probability distribution of force component errors and energy errors (relative to reference DFT calculations).

we find very good agreements between the GAP forces and DFT forces. The deviations of the atomic forces are generally within  $\pm 2$  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  $\pm 10$  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  $4.6 \text{ cm}^3/\text{mol}$  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 Section 3.1). Since the energies are sufficiently close between GAP and DFT, through the one-step thermodynamic perturbation method,  $\Delta F = k_B T \ln \langle e^{-\Delta U/k_B T} \rangle$ , 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 simula-

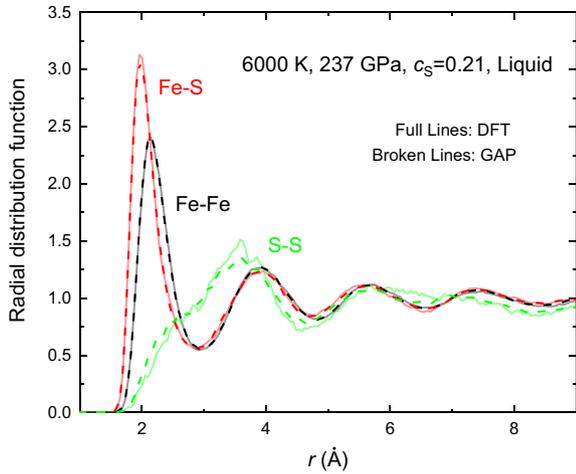
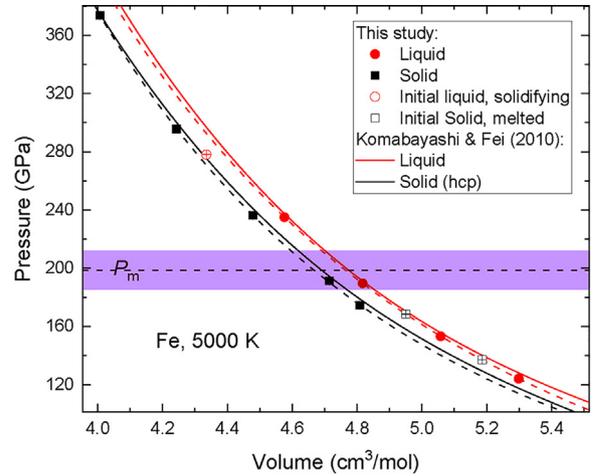


Fig. 2. Comparisons of the radial distribution functions of a Fe-S liquid (6000 K, 237 GPa and  $c_S = 0.21$ ) from the DFT simulation (full lines) and that with the GAP models (broken lines).

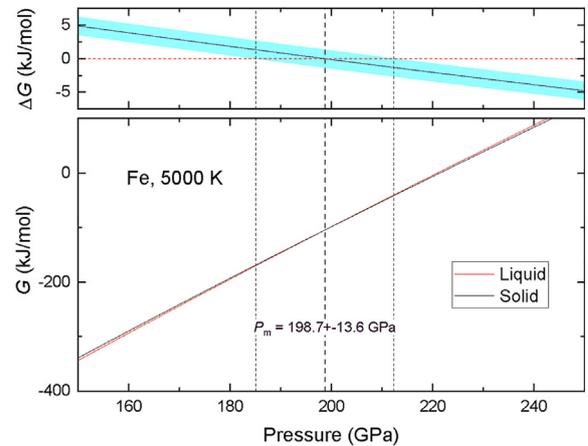
tion with GAP models has perfectly sampled the local structures around different atoms as compared with those from the DFT simulation. Just as previous findings from first principles simulations (Alfe and Gillan, 1998; Alfe et al., 2003), we can easily observe the net S-S repulsions and Fe-S attractions 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, mean square displacements and fluctuations of energies and pressures in the trajectories to verify the stabilities of the simulations within the simulation length scales. Based on these points we get the auxiliary Birch-Murnaghan equation of state for the liquid and solid respectively, which are shown in Fig. 3(a) with dashed curves and the numerical data can be found in Appendix C. It is apparent that the melted solids quickly establish new equilibrium within the duration of the simulation and their  $P$ - $V$  relations match the predictions of the liquid EOS quite well, while the solidifying liquids take much longer time to be fully crystalized and the averaged volumes/pressures are still slightly larger than the solid EOS predictions. These points give us rough estimates of the upper and lower limits of the melting pressure. In Fig. 3(a), we also include the curves provided by Komabayashi and Fei (2010) based on existing



(a)



(b)

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

experimental data and find very good agreements with the simulated EOS.

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

Table 2  
Simulated Helmholtz free energies ( $F$ ) of pure iron.

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

lower bound of the crossover by counting the errors of free energies.

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

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

From this equation and considering the error bars of the four points in Fig. 4, the melting temperature at 330 GPa is estimated to be  $6253 \pm 170$  K, comparing very well with  $6170 \pm 200$  K as recently estimated by Sun et al. (2018) and  $6350 \pm 300$  K as predicted by Alfe et al. (2002c). In contrast, simulations by Belonoshko et al. (2000) and Laio et al. (2000) give significantly higher (7100 K) or lower (5400 K) melting temperature, respectively, at the inner core boundary. With careful evaluations of the free energy errors from the classical potentials used in their studies, it is possible to correct these results to be in much better accordance with the DFT simulations (Alfe et al., 2002b).

Compared with the experiments, it is remarkable that our results almost perfectly agree with the measurements and extrapolations by Anzellini et al. (2013) over the entire pressure range. Shock wave measurements by Nguyen and Holmes (2004) and Brown and McQueen (1986) also fall into the same trend. On the other hand, diamond anvil

measurements by Boehler (1993) and Sinmyo et al (2019) and shock wave measurements by Yoo et al. (1993) give much lower or higher melting temperatures by up to over 1000 K. The exact reasons for these remarkable observed differences in experiments are still under debate due to the extreme technical challenges (Aquilanti et al., 2015; Morard et al., 2018).

### 4.3. Free energies of Fe-S alloys

For the Fe-S alloys, we deployed simulations with 9–36 sulfur atoms in the 180-atom cells (corresponding to  $c_S = 0.05 \sim 0.20$ , which is within the stability regime of *hcp*-structured Fe-S solid solutions according to previous studies (Cote et al., 2008; Gavryushkin et al., 2016)) and for each concentration we calculated its free energies through the thermodynamic integration techniques at several  $T$ - $P$  conditions (as listed in Table 3), based on which the free energies over the pressures can be derived through the auxiliary Birch-Murnaghan equation of states at various temperatures, as listed 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<sup>3</sup>/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 ( $NN_{S-S}$ ), as shown in the bottom plot of Fig. 6(a) and demonstrated in Fig. 6(b) by the radial distribution functions.  $NN_{S-S}$  decreases from the initial quasi-random arrangement value (around 43 here for a 180-atom cell with 36 sulfur atoms, i.e.,  $NN_{S-S} = 0.5zN_Sc_S$ , 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.,

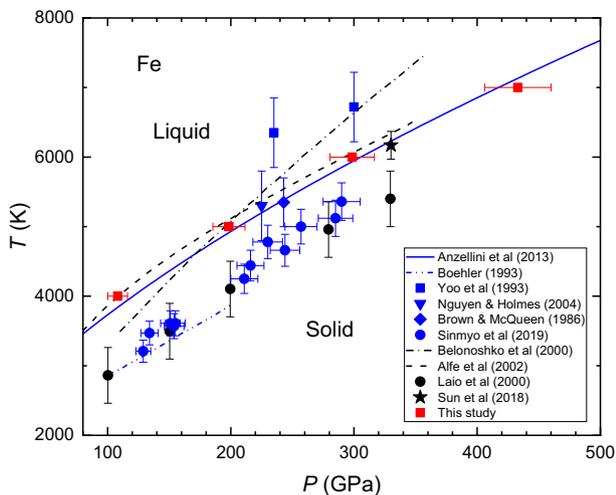


Fig. 4. Melting of iron from different approaches. Symbols and lines in blue are the experimental data, while those in black are the previous simulation results. The red squares are the melting points from this study with GAP models.

Table 3  
Simulated Helmholtz free energies ( $F$ ) of Fe-S alloys.

$T$ (K)	Phase	$c_S$	$V$ (cm <sup>3</sup> /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
		0.05	4.8086	189.94 ± 0.20	-1047.60 ± 0.92
5000	Liquid	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
		0.05	4.1225	329.36 ± 0.05	-878.78 ± 0.30
5000	Solid	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
		0.05	4.2395	329.51 ± 0.20	-1010.93 ± 1.67
6000	Liquid	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
		0.05	4.1636	329.27 ± 0.10	-987.78 ± 0.54
6000	Solid	0.10	4.1807	322.99 ± 0.10	-977.19 ± 0.43
		0.15	4.1692	325.74 ± 0.20	-954.92 ± 0.47

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

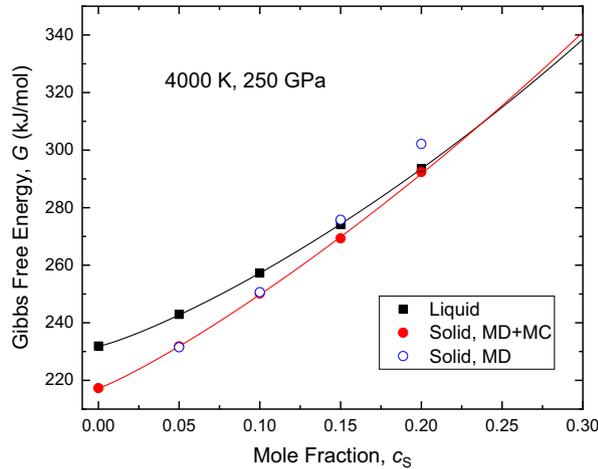
From the data points in Fig. 5, we obtained the free energy profiles over the composition range by linear least-square regressions with Eq. (10). The partitioning of sulfur in the coexisting liquid and solid iron phases can be firstly inferred from the variations of the two regressed parameters over temperatures and pressures. As listed in Table 4, the difference of  $\mu_S^{i,\text{sol}} - \mu_S^{i,\text{liq}}$  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  $\lambda_S^{\text{sol}} - \lambda_S^{\text{liq}}$  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 Eqs. (4) and (5), we quantitatively calculated the partition coefficients of sulfur ( $D_S = c_S^{\text{sol}}/c_S^{\text{liq}}$ ), as shown

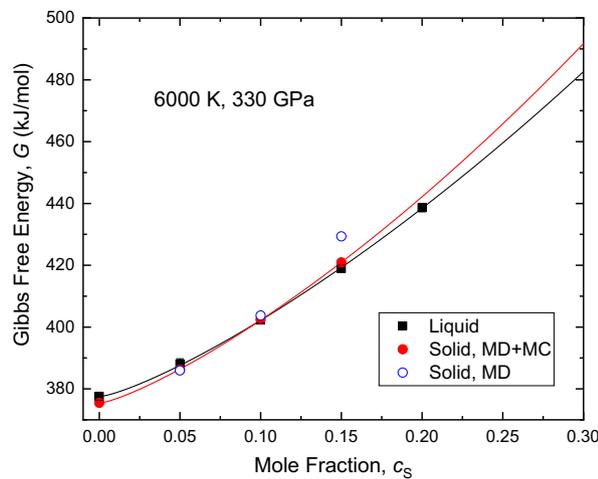
in Fig. 7. At 250 GPa and 330 GPa, similar values of  $D_S$  are predicted over temperature change of 1000 K, which reveals the minor temperature dependence of  $D_S$  and this is in accordance with the experimental findings (Kamada et al., 2010; Yokoo et al., 2019). For pressures lower than 250 GPa, 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 (Li et al., 2001; Stewart et al., 2007; Kamada et al., 2010; Kamada et al., 2012; Mori et al., 2017; 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  $0.75 \pm 0.07$ . This is in good agreement with the early prediction by Alfe et al. (2002a) through DFT simulations. The simple MD simulations with random solid solutions turn out to significantly underestimate  $D_S$  ( $\sim 0.53 \pm 0.04$ ) as illustrated in Fig. 7), which again emphasizes the importance of sufficient sampling with the hybrid Monte Carlo/molecular dynamics simulations.

## 5. IMPLICATIONS

According to our simulations in this study, the melting of pure iron at Earth's current inner core boundary pressure (330 GPa) results in a density jump ( $\Delta\rho \equiv \rho^{\text{sol}} - \rho^{\text{liq}}$ ) of 0.24 g/cm<sup>3</sup>, which amounts to 1.8% of  $\rho^{\text{sol}}$  (i.e.,  $\Delta\rho/\rho^{\text{sol}} \approx 1.8\%$ ) and compares very well with those predicted by previous simulations (as listed in the Table II of Sun et al. (2018)). This density jump is far smaller than the seismologically observed value of 0.6 g/cm<sup>3</sup> (Dziewonski and Anderson, 1981) or  $0.8 \pm 0.2$  g/cm<sup>3</sup> (Masters and Gubbins, 2003). Assuming Fe-S binary model for the Earth's core and the



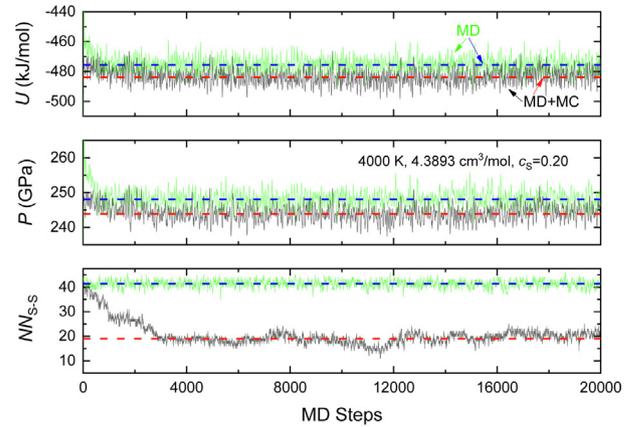
(a)



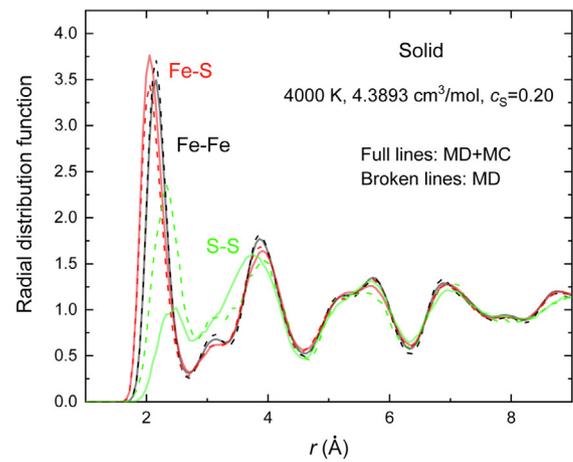
(b)

Fig. 5. Gibbs free energies of Fe-S solutions under 4000 K, 250 GPa (a) and 6000 K, 330 GPa (b). The filled black and red symbols are the free energies for liquids and solids, respectively. The lines are regressions with Eq. (10) based on these data (relevant parameters are listed in Table 4). Open blue circles are the results for solid solutions through simple molecular dynamics samplings, as compared with the filled red circles through hybrid Monte Carlo/molecular dynamics samplings (see the text). The uncertainties of the free energies are generally smaller than the sizes of the symbols.

temperature around 6000 K, we would need around 20% sulfur ( $c_S^{\text{liq}} \approx 0.20$ ) to match the outer core density ( $\rho^{\text{sol}} = 12.166 \text{ g/cm}^3$ , according to Dziewonski and Anderson (1981)) at ICB. With smaller amount of sulfur coexisting in the inner core,  $c_S^{\text{sol}} = D_S \times c_S^{\text{liq}} = 0.15$ , the density jump  $\Delta\rho$  increases to  $0.39 \text{ g/cm}^3$ . So the partitioning with sulfur alone is insufficient to account for the observed density jump across the Earth's inner core boundary and the Earth's core composition should be more complex than the simple Fe-S binary, which supports the findings of previous studies (Alfe et al., 2002a, 2003, 2007).



(a)



(b)

Fig. 6. Effects of Monte Carlo swapping on the samplings of thermodynamic properties and structures of the Fe-S solid solution. In (a), we show the propagations of total energy ( $U$ ), pressure ( $P$ ) and number of nearest-neighbor S-S pairs ( $NN_{S,S}$ ) during the trajectories of the two simulations (one with simple MD, in green; the other with hybrid Monte Carlo and MD, or MD + MC, in black) at 4000 K,  $4.3893 \text{ cm}^3/\text{mol}$ ,  $c_S = 0.20$ . The broken lines are equilibrium averaged values (i.e., discarding the first 5000 fs as pre-equilibrium stage) from both simulations (MD, in blue; MD + MC, in red) for better guides of the fluctuations. In (b), we compare the radial distribution functions of different pairs from the two simulations.

Although the exact core compositions are still elusive, geochemical studies often place around 2 wt% sulfur in the Earth's core (Allegre et al., 1995; McDonough, 2003; Wood et al., 2006). With this amount of sulfur, according to the comprehensive explorations by Badro et al (2014), the outer core would additionally need about 2.6 wt% oxygen and 1.8 wt% silicon to best fit the densities and seismic velocities at ICB and CMB (core mantle boundary). This leads to an outer core composition with  $c_O^{\text{liq}} = 0.083$ ,  $c_{\text{Si}}^{\text{liq}} = 0.033$  and  $c_S^{\text{liq}} = 0.032$ . Since  $D_{\text{Si}} \approx 1$ ,  $D_O < 0.01$  (calculated from the parameters provided by Alfe et al (2002a)) and  $D_S = 0.75$ , we would estimate the coexisting inner core

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

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

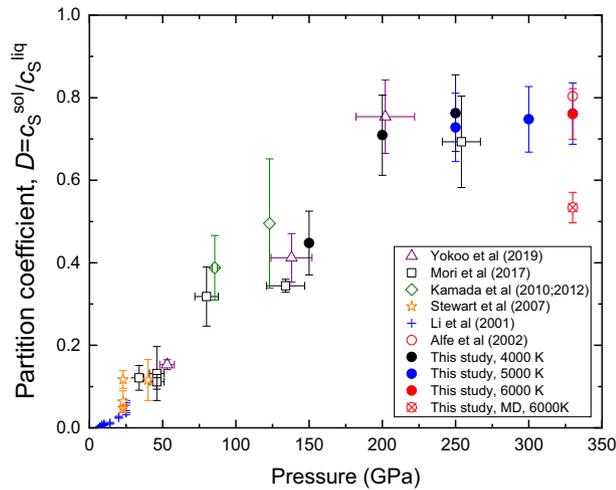


Fig. 7. Partition coefficients of sulfur between solid and liquid iron. Filled symbols are those from this study at different temperatures and pressures. The additional red crossed circle at 6000 K and 330 GPa is the result from simple MD simulations with random solid solutions. Open symbols are data from previous experiments (Kamada et al., 2012; Kamada et al., 2010; Li et al., 2001; Mori et al., 2017; Stewart et al., 2007; Yokoo et al., 2019) and simulations (Alfe et al., 2002a). Note that all the experimental data have been carefully converted to the ratio of mole fraction to be consistent with simulation results.

composition with about 1.7 wt% silicon and 1.4 wt% sulfur (i.e.,  $c_{\text{O}}^{\text{sol}} \approx 0$ ,  $c_{\text{Si}}^{\text{sol}} \approx c_{\text{Si}}^{\text{liq}} = 0.033$  and  $c_{\text{S}}^{\text{sol}} = D_{\text{S}} \times c_{\text{S}}^{\text{liq}} = 0.024$ ). Now with the solid solution model provided by Li et al (2018), the obtained inner core density would agree well with the PREM model by Dziewonski and Anderson (1981) with a relative error of 0.5% and the density jump would be around  $0.66 \text{ g/cm}^3$ . Note that we have not included hydrogen and carbon here since the properties of their alloys with iron are much more unclear, although they may be potentially important for explaining the seismic observations under certain circumstances (Li et al., 2018; Li et al., 2019; Mashino et al., 2019; Umemoto and Hirose, 2020).

If the above core composition is plausibly in accord with geochemical and geophysical constraints, then the roles of sulfur on the geodynamics of Earth's core should be minor. This may be inferred from its contribution to the density jump ( $0.66 \text{ g/cm}^3$  as mentioned above), which is important

for evaluating the gravitational energy to drive the geodynamo (Stacey and Stacey, 1999; Gubbins et al., 2004). By “turning on” the partitioning coefficients stepwise (i.e., we inspect the differences by switching the  $D_{\text{O}}$  and  $D_{\text{S}}$  from 1.0 to the expected values), the exclusion of oxygen in the inner core and partitioning of sulfur would account for about  $0.44 \text{ g/cm}^3$  and  $0.04 \text{ g/cm}^3$  respectively, in addition to  $0.18 \text{ g/cm}^3$  from the assumed congruent freezing (i.e., the solid is assumed to be in the same composition with the coexisting liquid). It seems that the 6% contribution here from sulfur cannot be entirely neglected, but this is likely to be an upper bound, since sulfur is recently found to be less siderophile in the Earth's core (Suer et al., 2017).

Finally, the almost invariant partition coefficient from 250 GPa to higher pressures is interesting. 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., multi-component 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 with fundamentally rigorous thermodynamic integration method within affordable computational cost. In fact, to thoroughly sample the phase space around the liquidus and solidus of Fe-S solutions, we have carried out over 500 independent atomistic simulations in this study, each with 180 atoms and at least 20,000 steps. The current implementation of the GAP models takes about 12 CPU seconds for each step in each run. With parallel acceleration of 24 CPU cores for each run, we have managed to accomplish the simulations within two months. As a comparison, the direct high precision DFT simulations are about three orders of magnitudes slower than the GAP simulations.

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

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

#### RESEARCH DATA FOR THIS ARTICLE

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

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### ACKNOWLEDGEMENTS

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

#### SUPPLEMENTARY MATERIALS

Supplementary materials to this article can be found in: Appendix A that includes several supporting figures for the main text; Appendix B that has the details for a GAP model training with the QUIP code; Appendix C that includes the complementary research data and goes to the Electronic Annex. These materials can be found online at <https://doi.org/10.1016/j.gca.2020.03.028>.

#### REFERENCES

- Alboussiere T., Deguen R. and Melzani M. (2010) Melting-induced stratification above the Earth's inner core due to convective translation. *Nature* **466**, 744–U749.
- Alfè D. (2009) PHON: A program to calculate phonons using the small displacement method. *Comput. Phys. Commun.* **180**, 2622–2633.
- Alfe D. and Gillan M. J. (1998) First-principles simulations of liquid Fe-S under Earth's core conditions. *Phys Rev B* **58**, 8248–8256.
- Alfe D., Gillan M. J. and Price G. D. (2002a) Ab initio chemical potentials of solid and liquid solutions and the chemistry of the Earth's core. *J. Chem. Phys.* **116**, 7127–7136.
- Alfe D., Gillan M. J. and Price G. D. (2002b) Complementary approaches to the ab initio calculation of melting properties. *J. Chem. Phys.* **116**, 6170–6177.
- Alfe D., Price G. D. and Gillan M. J. (2002c) Iron under Earth's core conditions: Liquid-state thermodynamics and high-pressure melting curve from ab initio calculations. *Phys. Rev. B* **65**.
- Alfe D., Gillan M. J. and Price G. D. (2003) Thermodynamics from first principles: temperature and composition of the Earth's core. *Mineral. Mag.* **67**, 113–123.
- Alfe D., Gillan M. J. and Price G. D. (2007) Temperature and composition of the Earth's core. *Contemp. Phys.* **48**, 63–80.
- Allegre C. J., Poirier J. P., Humler E. and Hofmann A. W. (1995) The chemical-composition of the earth. *Earth Planet. Sc. Lett.* **134**, 515–526.
- Anzellini S., Dewaele A., Mezouar M., Loubeyre P. and Morard G. (2013) Melting of iron at Earth's inner core boundary based on fast X-ray diffraction. *Science* **340**, 464–466.
- Aquilanti G., Trapananti A., Karandikar A., Kantor I., Marini C., Mathon O., Pascarelli S. and Boehler R. (2015) Melting of iron determined by X-ray absorption spectroscopy to 100 GPa. *P Natl. Acad. Sci. USA* **112**, 12042–12045.

- Arfken G. (1985) Stirling's Series. In *Mathematical methods for physicists* (ed. F. Orlando), 3rd ed. Academic Press, pp. 555–559.
- Badro J., Côté A. S. and Brodholt J. P. (2014) A seismologically consistent compositional model of Earth's core. *Proc. Natl. Acad. Sci.* **111**, 7542–7545.
- Bartók A. P., Kermode J., Bernstein N. and Csányi G. (2018) Machine learning a general-purpose interatomic potential for silicon. *Phys. Rev. X* **8**, 041048.
- Bartók A. P., Gillan M. J., Manby F. R. and Csányi G. (2013) Machine-learning approach for one- and two-body corrections to density functional theory: Applications to molecular and condensed water. *Phys. Rev. B* **88**, 054104.
- Bartók A. P., Kondor R. and Csányi G. (2013) On representing chemical environments. *Phys. Rev. B* **87**, 184115.
- Bartók A. P., Payne M. C., Kondor R. and Csányi G. (2010) Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* **104**, 136403.
- Behler J. and Parrinello M. (2007) Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **98**, 146401.
- Belonoshko A. B., Ahuja R. and Johansson B. (2000) Quasi - Ab initio molecular dynamic study of Fe melting. *Phys. Rev. Lett.* **84**, 3638–3641.
- Ben-Naim A. (2006) *Molecular theory of solutions*. Oxford University Press, New York.
- Birch F. (1964) Density and composition of mantle and core. *J. Geophys. Res.* **69**, 4377–4388.
- Boehler R. (1993) Temperatures in the Earth's core from melting-point measurements of iron at high static pressures. *Nature* **363**, 534–536.
- Brown J. M. and McQueen R. G. (1986) Phase transition, Grüneisen parameter, and elasticity for shocked iron between 77 GPa and 400 GPa. *J. Geophys. Res.* **91**, 7485–7494.
- Cao A. and Romanowicz B. (2004) Constraints on density and shear velocity contrasts at the inner core boundary. *Geophys. J. Int.* **157**, 1146–1151.
- Cerriotti M., Willatt M. J. and Csányi G. (2018) Machine learning of atomic-scale properties based on physical principles. In *Handbook of Materials Modeling: Methods: Theory and Modeling* (eds. W. Andreoni and S. Yip). Springer International Publishing, Cham, pp. 1–27.
- Chen B., Li J. and Hauck S. A. (2008) Non-ideal liquidus curve in the Fe-S system and Mercury's snowing core. *Geophys. Res. Lett.* **35**, L07201.
- Chudinovskikh L. and Boehler R. (2007) Eutectic melting in the system Fe–S to 44 GPa. *Earth Planet. Sc. Lett.* **257**, 97–103.
- Cote A. S., Vocadlo L. and Brodholt J. P. (2008) Light elements in the core: Effects of impurities on the phase diagram of iron. *Geophys. Res. Lett.* **35**, L05306.
- Cottaar S. and Buffett B. (2012) Convection in the Earth's inner core. *Phys Earth Planet In* **198**, 67–78.
- Deguen R. and Cardin P. (2011) Thermochemical convection in Earth's inner core. *Geophys. J. Int.* **187**, 1101–1118.
- Deringer V. L., Caro M. A. and Csányi G. (2019) Machine learning interatomic potentials as emerging tools for materials science. *Adv. Mater.*, 1902765.
- Deringer V. L., Csányi G. and Proserpio D. M. (2017) Extracting crystal chemistry from amorphous carbon structures. *Chem-PhysChem* **18**, 873–877.
- Dorner F., Sukurma Z., Dellago C. and Kresse G. (2018) Melting Si: beyond density functional theory. *Phys. Rev. Lett.* **121**, 195701.
- Dragoni D., Daff T. D., Csányi G. and Marzari N. (2018) Achieving DFT accuracy with a machine-learning interatomic potential: Thermomechanics and defects in bcc ferromagnetic iron. *Phys. Rev. Mater.* **2**, 013808.
- Drautz R. (2019) Atomic cluster expansion for accurate and transferable interatomic potentials. *Phys. Rev. B* **99**, 014104.
- Dziewonski A. M. and Anderson D. L. (1981) Preliminary reference Earth model. *Phys. Earth Planet Inter.* **25**, 297–356.
- Gavryushkin P. N., Popov Z. I., Litasov K. D., Belonoshko A. B. and Gavryushkin A. (2016) Stability of B2-type FeS at Earth's inner core pressures. *Geophys. Res. Lett.* **43**, 8435–8440.
- Gubbins D., Alfe D., Masters G., Price G. D. and Gillan M. (2004) Gross thermodynamics of two-component core convection. *Geophys. J. Int.* **157**, 1407–1414.
- Gubbins D., Alfe D. and Davies C. J. (2013) Compositional instability of Earth's solid inner core. *Geophys. Res. Lett.* **40**, 1084–1088.
- Hirose K., Labrosse S. and Hernlund J. (2013) Composition and State of the Core. *Annu. Rev. Earth Planet. Sci.* **41**(41), 657–691.
- Kamada S., Ohtani E., Terasaki H., Sakai T., Miyahara M., Ohishi Y. and Hirao N. (2012) Melting relationships in the Fe-Fe3S system up to the outer core conditions. *Earth Planet. Sc. Lett.* **359**, 26–33.
- Kamada S., Terasaki H., Ohtani E., Sakai T., Kikegawa T., Ohishi Y., Hirao N., Sata N. and Kondo T. (2010) Phase relationships of the Fe-FeS system in conditions up to the Earth's outer core. *Earth Planet. Sc. Lett.* **294**, 94–100.
- Komabayashi T. and Fei Y. W. (2010) Internally consistent thermodynamic database for iron to the Earth's core conditions. *J. Geophys. Res.-Sol. Ea* **115**, B03202.
- Kresse G. and Joubert D. (1999) From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775.
- Kung S. Y. (2014) *Kernel Methods and Machine Learning*. Cambridge University Press, Cambridge.
- Labrosse S. (2014) Thermal and compositional stratification of the inner core. *C.R. Geosci.* **346**, 119–129.
- Laio A., Bernard S., Chiarotti G. L., Scandolo S. and Tosatti E. (2000) Physics of iron at Earth's core conditions. *Science* **287**, 1027–1030.
- Li J. and Fei Y. (2007) 2.14 - Experimental Constraints on Core Composition. In *Treatise on Geochemistry* (eds. H. D. Holland and K. K. Turekian). Pergamon, Oxford, pp. 1–31.
- Li J., Fei Y., Mao H. K., Hirose K. and Shieh S. R. (2001) Sulfur in the Earth's inner core. *Earth Planet. Sc. Lett.* **193**, 509–514.
- Li Y., Vočadlo L. and Brodholt J. P. (2018) The elastic properties of hcp-Fe alloys under the conditions of the Earth's inner core. *Earth Planet. Sc. Lett.* **493**, 118–127.
- Li Y., Vocadlo L., Alfe D. and Brodholt J. (2019) Carbon partitioning between the Earth's inner and outer core. *J. Geophys. Res.-Sol. Ea* **124**, 12812–12824.
- Lythgoe K. H., Rudge J. F., Neufeld J. A. and Deuss A. (2015) The feasibility of thermal and compositional convection in Earth's inner core. *Geophys. J. Int.* **201**, 764–782.
- Ma Z. T. (2001) Thermodynamic description for concentrated metallic solutions using interaction parameters. *Metall. Mater. Trans. B* **32**, 87–103.
- Mashino I., Miozzi F., Hirose K., Morard G. and Sinmyo R. (2019) Melting experiments on the Fe-C binary system up to 255 GPa: Constraints on the carbon content in the Earth's core. *Earth Planet. Sc. Lett.* **515**, 135–144.
- Masters G. and Gubbins D. (2003) On the resolution of density within the Earth. *Phys. Earth Planet. In* **140**, 159–167.
- McDonough W. F. (2003) Compositional Model for the Earth's Core. In *Treatise on Geochemistry* (eds. H. D. Holland and K. K. Turekian). Pergamon, Oxford, pp. 547–568.

- McDonough W. F. and Sun S. S. (1995) The composition of the Earth. *Chem. Geol.* **120**, 223–253.
- Mermin N. D. (1965) Thermal properties of inhomogeneous electron Gas. *Phys. Rev.* **137**, 1441–1443.
- Morard G., Andrault D., Antonangeli D. and Bouchet J. (2014) Properties of iron alloys under the Earth's core conditions. *C.R. Geosci.* **346**, 130–139.
- Morard G., Andrault D., Guignot N., Sanloup C., Mezouar M., Petitgirard S. and Fiquet G. (2008) In situ determination of Fe-Fe<sub>3</sub>S phase diagram and liquid structural properties up to 65 GPa. *Earth Planet. Sc. Lett.* **272**, 620–626.
- Morard G., Boccato S., Rosa A. D., Anzellini S., Miozzi F., Henry L., Garbarino G., Mezouar M., Harmand M., Guyot F., Boulard E., Kantor I., Irifune T. and Torchio R. (2018) Solving controversies on the iron phase diagram under high pressure. *Geophys. Res. Lett.* **45**, 11074–11082.
- Mori Y., Ozawa H., Hirose K., Sinmyo R., Tateno S., Morard G. and Ohishi Y. (2017) Melting experiments on Fe-Fe<sub>3</sub>S system to 254 GPa. *Earth Planet. Sc. Lett.* **464**, 135–141.
- Nguyen J. H. and Holmes N. C. (2004) Melting of iron at the physical conditions of the Earth's core. *Nature* **427**, 339–342.
- Perdew J. P., Burke K. and Ernzerhof M. (1996) Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868.
- Poirier J.-P. (1994) Light elements in the Earth's outer core: A critical review. *Phys. Earth Planet. In* **85**, 319–337.
- Rupp M. (2015) Special issue on machine learning and quantum mechanics. *Int. J. Quantum. Chem.* **115**, 1003–1004.
- Shapeev A. (2017) Accurate representation of formation energies of crystalline alloys with many components. *Comp. Mater. Sci.* **139**, 26–30.
- Sinmyo R., Hirose K. and Ohishi Y. (2019) Melting curve of iron to 290 GPa determined in a resistance-heated diamond-anvil cell. *Earth Planet. Sc. Lett.* **510**, 45–52.
- Stacey F. D. and Stacey C. H. B. (1999) Gravitational energy of core evolution: implications for thermal history and geodynamo power. *Phys. Earth Planet. In.* **110**, 83–93.
- Stewart A. J., Schmidt M. W., van Westrenen W. and Liebske C. (2007) Mars: A new core-crystallization regime. *Science* **316**, 1323–1325.
- Suer T. A., Siebert J., Remusat L., Menguy N. and Fiquet G. (2017) A sulfur-poor terrestrial core inferred from metal-silicate partitioning experiments. *Earth Planet. Sc. Lett.* **469**, 84–97.
- Sun T., Brodholt J. P., Li Y. G. and Vocadlo L. (2018) Melting properties from ab initio free energy calculations: Iron at the Earth's inner-core boundary. *Phys. Rev. B* **98**, 224301.
- Szlachta W. J., Bartok A. P. and Csanyi G. (2014) Accuracy and transferability of Gaussian approximation potential models for tungsten. *Phys. Rev. B* **90**, 104108.
- Terasaki H., Kamada S., Sakai T., Ohtani E., Hirao N. and Ohishi Y. (2011) Liquidus and solidus temperatures of a Fe-O-S alloy up to the pressures of the outer core: Implication for the thermal structure of the Earth's core. *Earth Planet. Sc. Lett.* **304**, 559–564.
- Umamoto K. and Hirose K. (2020) Chemical compositions of the outer core examined by first principles calculations. *Earth Planet. Sc. Lett.* **531**(8), 116009.
- Veit M., Jain S. K., Bonakala S., Rudra I., Hohl D. and Csanyi G. (2019) Equation of state of fluid methane from first principles with machine learning potentials. *J. Chem. Theory Comput.* **15**, 2574–2586.
- White, V.W.M., 2013. *Geochemistry*. Wiley-Blackwell.
- Widom M., Huhn W. P., Maiti S. and Steurer W. (2014) Hybrid Monte Carlo/molecular dynamics simulation of a refractory metal high entropy alloy. *Metall. Mater. Trans. A* **45A**, 196–200.
- Wood B. J., Walter M. J. and Wade J. (2006) Accretion of the Earth and segregation of its core. *Nature* **441**, 825–833.
- Yokoo S., Hirose K., Sinmyo R. and Tagawa S. (2019) Melting experiments on liquidus phase relations in the Fe-S-O ternary system under core pressures. *Geophys. Res. Lett.* **46**, 5137–5145.
- Yoo C. S., Holmes N. C., Ross M., Webb D. J. and Pike C. (1993) Shock temperatures and melting of iron at earth core conditions. *Phys. Rev. Lett.* **70**, 3931–3934.

Associate editor: Marc Blanchard