A grand canonical approach for modelling hydrogen trapping at vacancies in α-Fe

E. R. M. Davidson,[1] T. Daff,[2] G. Csanyi,[2] and M. W. Finnis[1]

[1]*Department of Materials and Department of Physics, Thomas Young Centre, Imperial College London, Exhibition Road, SW7 2AZ London, United Kingdom*

[2]*Engineering Laboratory, University of Cambridge, Cambridge CB2 1PZ, United Kingdom*

Abstract

Vacancies in iron are hydrogen traps, important in the understanding of hydrogen embrittlement of steel. We present a grand canonical approach to computing the trap occupancy as a function of both temperature and hydrogen concentration from practically zero to super-saturation. Our method couples a purpose-made machine-learned H-Fe potential, which enables rapid sampling with near density functional theory accuracy, with a statistical mechanical calculation of the trap occupancy using the technique of nested sampling. In contrast to the conventional assumption (based on Oriani theory) that at industrially relevant hydrogen concentrations and ambient conditions vacancy traps are are fully occupied, we find that vacancy traps are less than fully occupied under these conditions, necessitating a reevaluation of how we think about "mobile hydrogen" in iron and steel.

I. INTRODUCTION

The propensity of hydrogen to embrittle steels is an important practical problem, motivating physicists, electro-chemists and corrosion engineers to understand the absorption, trapping and diffusion of hydrogen, its devastating effect on ductility, and how it can be controlled. Although a number of mechanisms for embrittlement have been put forward, their relative importance is still a matter of ongoing research and debate [1–5]. Whatever the particular mechanism of embrittlement, hydrogen traps, including precipitates, dislocations and vacancies, may play a protective role by sequestering the hydrogen [6]. The possibility of this has recently been observed by atomic probe microscopy of deuterium in carbide particles [7]. Static calculations have strongly suggested that a single vacancy in α-Fe can trap up to 6 H atoms [8–10], although the marginal trapping energy of the 6th is very small. The maximum number was challenged by Sugimoto and Fukai [11], who concluded from their Monte-Carlo simulations that it should be 5, a conclusion supported by an earlier DFT calculation by Tateyama and Ohno [12]. The lowering of vacancy formation energy by hydrogen trapping is the origin of the superabundant vacancy effect discovered by Fukai and Ōkuma [13]. Trapping by vacancies and the superabundant vacancy effect have been investigated by many authors since, for a wide range of metals [14–22]. This formation energy lowering was described by Kirchhheim [15, 16] as a 'defactant' effect, by analogy with Gibbs' original theory of surfactants. The defactant effect is a step in at least one of the

proposed embrittlement mechanisms, sometimes referred to as Hydrogen Enhanced Stress Induced Vacancy Formation (HESIV) [1, 23]. We know that the solubility of hydrogen in iron is very low [24], but that under operating conditions it only requires less than one ppm dissolved in the crystalline matrix of a steel to cause embrittlement. In the present paper we are addressing the question of how many hydrogens would be sequestered in a vacancy under realistic conditions. The point here is not whether the maximum number at low temperature is 5 or 6, but how the mean number trapped varies with temperature and the bulk concentration or, equivalently, chemical potential.

A simple "lattice model" was proposed by Oriani [25], which represents the bcc Fe by discrete tetrahedral interstitial sites and the vacancy by six trap sites, corresponding to the static equilibrium locations of the H, which in the vacancy are close to sites that in the perfect crystal are octahedral. The average occupancy of a vacancy can then be expressed as

$$\langle n_v \rangle = \frac{6\, y\, e^{-\beta \Delta E}}{1 + y\, e^{-\beta \Delta E}}, \tag{1}$$

where y is the fractional lattice concentration, $\beta = 1/(k_B T)$, and ΔE is the binding energy of the vacancy trap relative to a lattice site. More sophisticated lattice models, which include the distinct configurations of multiple hydrogens within a vacancy and the interactions between them have been developed, a good example for our specific case of bcc Fe is the work of Tateyama and Ohno [12]. In the case of Ni, the vibrational contribution to the free energy of the atomic nuclei has been included within a quasiharmonic approximation [21] but the major limitation of a lattice model remains, namely that the H atoms are confined to discrete sites. To the best of our knowledge, the only model that also included the full contribution to the free energy from off-lattice movement of the hydrogen is that of Tanguy and Mareschal [14], and Tanguy *et al.* [26], for Al and Ni respectively. For Al, they employed an Off-Lattice Monte-Carlo (OLMC) approach, with an EAM model of the Al-H interaction to evaluate the superabundant vacancy effect. For Ni they also made detailed calculations with DFT for H clusters in vacancies, where H was located on the various high-symmetry sites, but used OLMC and and an EAM potential to estimate the off-lattice contribution.

For the case of interest here, the weak binding energy ($< 0.1\,\mathrm{eV}$) [27, 28] to the lowest energy, tetrahedral sites in Fe suggests that in reality H moves smoothly and swiftly in a continuous landscape of potential, even at room temperature, implying that its total entropy in the bulk will be significantly greater than the configurational entropy captured by the

lattice model. For the same reason its entropy when trapped in a vacancy will also be greater, so the net effect of the lattice approximation on trap occupancy is unclear. How to calculate the hydrogen occupancy of a trap using a simple yet accurate approach beyond the lattice model is the challenge that we address here.

Density functional theory [29, 30] is the most popular first-principles method for computing the energies and energy barriers, due to its good compromise between computational cost and accuracy. However, for sampling configurational phase space including a vacancy, it is too computationally demanding. Empirical interatomic potentials are much faster, but their accuracy and transferability are always questionable. A possible solution to this problem is to create a moderately large dataset of DFT calculations on configurations strictly relevant to the problem at hand, and then fit a potential with a very flexible functional form, which can reliably predict the energetics of the H-Fe system including a vacancy, but without necessarily being transferable to *very* different configurations, such as surfaces. Recently various neural network, Gaussian process regression and similar approaches have been applied to create such potentials [31–34]. Here we demonstrate the first Gaussian Approximation Potential (GAP) for the H – α-Fe interaction.

The statistical mechanical solution to the problem of trap occupancy can be formulated in terms of a grand canonical ensemble and a partition function, the calculation of which requires sampling the potential energy surface, weighted by a Boltzmann factor. For our scenarios in which the bulk concentration is very low, we can conveniently express the result in terms of readily-evaluated canonical partition functions, as we show below.

A number of methods have become widely used to enable the efficient sampling of potential energy surfaces, including metadynamics [35], Wang-Landau sampling [36], parallel tempering [37, 38] and nested sampling [39, 40]. We have chosen the latter approach, because it allows the direct calculation of the partition function itself, which is essential to allow access to the regime of very low H concentrations. The problem with low H concentrations arises due to the *supercell* approximation (a small periodically repeating block of unit cell of crystal containing a single vacancy). The very low probability of finding a H atom anywhere within a supercell which allows accurate DFT-quality calculations is very low, leading to very poor sampling of the relative probability of the H being trapped or not. But we can

turn the low H concentration to our advantage, by sampling separately the configuration space of a *single* H atom within a relatively small supercell of *perfect* crystal, which enables us to calculate the chemical potential of an arbitrarily dilute, ideal solution of H in Fe. This chemical potential is then used as the formal boundary condition for the *grand* canonical sampling of H trapped by the vacancy, which is thus insensitive to the definition of the supercell that contains the vacancy. Since our approach considers both the full phase space of the bulk and vacancy regions of the system and includes H-H interactions, we obtain far more realistic results than the lattice model, and can comment on the effects of the approximations made in the Oriani approach. This new combination of a machine-learned potential with nested sampling and the grand canonical ensemble should be useful for analysing other trapping problems.

II. MODELS FOR CALCULATING ENERGIES AND FORCES

We model the system using a 54-atom cubic supercell ($3 \times 3 \times 3$ body-centred unit cells), with periodic boundary conditions. It is well known that in bulk Fe, the four-fold coordinated tetrahedral sites are the preferred sites for interstitial H-atoms [27, 28], and we have confirmed this, noting however, that they are very shallow traps, separated by a barrier of only about 88 meV. Figure 1 shows the potential energy surface of H in the presence of a Fe vacancy, with six equivalent trap positions, close to the location of the cube face-centers of the body-centred unit cell (octahedral sites). In all the calculations in this work, the positions of the Fe atoms are fixed after an initial DFT relaxation of the vacancy, and so there is no need to model the Fe subsystem, only its interaction with H. We now give a brief overview of the Gaussian Approximation Potential methodology that we use. A more extensive discussion is found in papers reporting recent applications [32, 41–46] and a review [47]. The key idea is to posit interatomic functional forms that are highly flexible due to a large number of parameters (universal, in the limit), but avoid overfitting because the least squares problem that we solve to obtain the parameters is highly regularised, resulting in smooth potential functions. This still leaves a lot of freedom for designing potentials, in particular in the choice of representation of the atomic geometry. We find that the customary decomposition of the total energy into two-body, three-body, etc. terms is very helpful. The two-body functions are parameterized by the interatomic distance, whereas

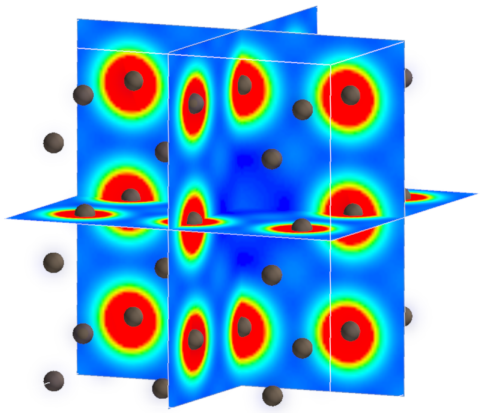

FIG. 1. Representation of 53-atom $(3 \times 3 \times 3)$ supercell with a vacancy at the centre. The potential energy for a single H-atom in the cell is represented by the heat maps on the cross sections. The six trap sites correspond to the dark-blue regions.

the three-body terms are described by a symmetrized combination of the three distances. For the basis functions, we choose Gaussians, because they achieve our dual goals of being universal approximators and are intuitive to parametrise. This formalism was successfully used recently for the low body-order terms of potentials for amorphous carbon [46] and boron [45]. Whereas in those cases, a many-body term was also needed, we found that in the present case, just two- and three-body terms are sufficient. We found it sufficient to include explicit three-body terms for the interaction of an H atom with two Fe neighbours, and there are no explicit three-body terms for the interaction of two H atoms and a Fe atom or three H atoms.

The interaction between pairs of H-atoms, which reduces the magnitude of the trap energy as the vacancy fills, is described by a Yukawa potential with a screening length of 1.5 Å^{-1}. The interaction energy between a H atom and a collection of Fe atoms is given by

$$E_{\mathrm{H-Fe}}(r_{\mathrm{H}}) = \sum_{i \in \{\mathrm{Fe}\}} V_2(|r_H - r_i|) + \sum_{j<k \in \{\mathrm{Fe}\}} V_3(r_H, r_j, r_k). \tag{2}$$

Within the GAP framework, the two body potential V_2 is written as a linear combination of kernel basis functions,

$$V_2(r) = \sum_m x_{2,m} K_2(r, r_m), \tag{3}$$

where the coefficients $x_{2,m}$ are to be fitted, and we choose a Gaussian kernel $K_2(r, r_m) = \exp(-|r - r_m|^2/2\theta_2^2)$ with width parameter $\theta_2 = 1.5$ Å. The potential is smoothly taken to

zero at $r = 6$ Å.

The three body term, V_3, describes the interaction of a H atom (i) with two Fe neighbours (j and k), and is constructed similarly to the two-body potential, except that its arguments are explicitly made symmetric with respect to swapping Fe atoms. The new symmetrized coordinates are

$$q_1 = r_{ij} + r_{ik} \tag{4}$$

$$q_2 = (r_{ij} - r_{ik})^2 \tag{5}$$

$$q_3 = r_{jk}, \tag{6}$$

collected into a vector $\mathbf{q} = (q_1, q_2, q_3)$, and so the symmetrized three-body term $\tilde{V}_3$ is given by

$$V_3(r_i, r_j, r_k) \equiv \tilde{V}(\mathbf{q}) = \sum_m x_{3,m} K_3(\mathbf{q}, \mathbf{q}_m), \tag{7}$$

where again the coefficients $x_{3,m}$ are to be fitted, K_3 is a three dimensional Gaussian with three independent length scale parameters, which were set to be 1/2 of the extent of the training data in each dimension. The cutoff in this case was 5 Å. The training data for the H-Fe interaction potential comprises snapshots from molecular dynamics trajectories of 54 and 128 Fe atoms with either 0, 1, or 2 Fe atoms removed, and a single H atom added. Plane wave DFT calculations were performed for each configuration with and without the H atom, and the interaction energy defined as the difference between the two total energies. All parameter settings of the DFT calculations were identical to those reported for pure Fe calculations by Dragoni *et al.* [48]. Altogether, about 400 configurations were used in the fit, comprising about 28k atoms. The fit is obtained as the analytic solution to a linear least-squares problem, where the target data is the H-Fe interaction energy and corresponding gradient. The GAP framework uses Tikhonov regularisation, we set the magnitude to 10^{-5} eV and 10^{-4} eV/Å for the energy and forces, respectively. The mean absolute errors of the fit on an independent test set were 20 meV for energies and 10 meV/Å for force components. For comparison, the mean absolute errors of the EAM model of Ref. [49] for the H-Fe interaction on the same set of configurations are 250 meV for energies and 730 meV/Å for force components. The resulting potentials for the two and three body interactions energies are shown in Fig. 2. We used the QUIP software package [50] to obtain the fit and its PYTHON API to subsequently calculate the potential and its gradients. More details and

numerical values of the parameters for these potentials can be downloaded [51]. Employing

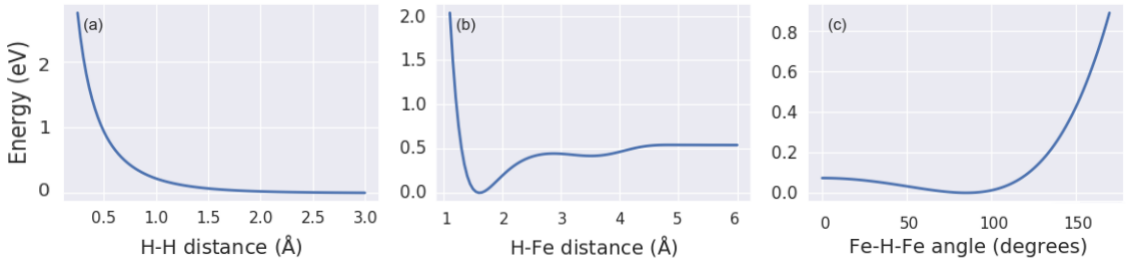

FIG. 2. The two and three body interaction energies used in this work. The Yukawa potential for the H-H interaction energy is shown in (a), (b) shows the 2-body GAP H-Fe potential, and (c) shows the 3-body GAP Fe-H-Fe potential upon varying the angle.

the interatomic potential enables us to compute the energy of a configuration at nearly DFT accuracy in around 100 ms on a single CPU core using our naive implementation of the 2- and 3-body functions. The same calculation would take around 15 minutes on 24 cores of the current UK national supercomputer. This already represents an enormous saving in terms of both time and cost, and was critical in enabling us to perform nested sampling calculations. The above potential functions could easily be tabulated and used within a software package such as LAMMPS thereby obtaining significant further increases in speed.

III. SAMPLING

Our chosen method for the statistical mechanics sampling is *nested sampling*, originally developed by Skilling to estimate the Bayesian evidence for model comparison purposes [39, 40]. The approach was adapted by Pártay *et al.* to compute directly and efficiently the partition function in atomistic problems, and it enables us to compute any thermodynamic property as a function of temperature from a single run [52]. Briefly, a set of configurations x_n of decreasing energy $E_n = E(x_n)$ is generated with the property that the ratio of phase space volumes enclosed by successive energy level sets E_n is approximately constant, $\alpha = K/(K+1)$, where K is the number of samples, held constant. We start by chosing K configurations randomly, and in iteration n, the highest energy sample is assigned as x_n, and is replaced by a new sample that is randomly chosen from among the others, cloned, and then decorrelated by a bounded random walk in which we reject any move that would result

in an energy greater than E_n. At the end, the canonical partition function is estimated (up to a multiplicative constant) as,

$$Q_C(T) \simeq \sum_n (\alpha^n - \alpha^{n+1}) e^{-\beta E(x_n)}. \tag{8}$$

To test convergence with respect to the number of walkers, K, we ran four independent copies of each simulation. We found that 192 walkers were sufficient.

All our nested sampling calculations were carried out with the PYMATNEST package[53].

IV. EVALUATING THE GRAND CANONICAL PARTITION FUNCTION

Despite the efficiency gains introduced by the methods above, direct calculation of thermodynamic properties in the canonical ensemble at the low concentrations ($< 100\,\text{ppm}$) of H that are damaging in real systems would be impractical due to the large size of the simulation cell required ($> 10^4$ atoms). We can nevertheless compute the expected number of H-atoms in the vacancy at any concentration by conceptually separating the vacancy region from the bulk region, and treating the former with variable number of H atoms within a grand canonical ensemble. The phrase "in the vacancy" implies that we can define a boundary to the vacant lattice site. The position of this boundary is in principle arbitrary, but for practical purposes, as long as the bulk concentration is sufficiently low, the number of H atoms deemed to be inside the vacancy will not be significantly changed by the choice of this boundary within reasonable limits. This can easily be verified in the case considered in this paper, for which the borders of a $3 \times 3 \times 3$ supercell are a suitable boundary. It contains 54 atomic sites. Suppose the bulk concentration of H were 10^{-4} per atom, which is the largest value we consider. Assuming a Poisson distribution it is easy to show that less than 0.5% of such bulk supercells would be occupied by any hydrogen, and less than 0.3% of those that are would be occupied by more than one atom of hydrogen. This leads us to consider the H content of a bulk supercell compared to one of the same size containing a vacancy, with which the hydrogen is in equilibrium. If the bulk concentration is only significantly perturbed at the atoms neighbouring the vacant site, then the mean hydrogen content of the supercell $\langle n_v \rangle$ would be equal to the excess hydrogen in the vacancy to within 0.005 atoms, which is more than adequate for our purposes. Longer ranged strain-fields will introduce a small error in our estimates, which would be of minor importance in practice. Thus it is

unnecessary besides physically meaningless to define a bounding surface any closer to the vacant site than the borders of the supercell. This is further justified in the Appendix.

To derive the required formula for computation, we adopt from the start a notation which defines discrete states in energy or space, labeled with subscript i, summations over which can be equated to integrals weighted by a density of states. The expectation value for the number of H-atoms n_v in the vacancy in the grand canonical ensemble can be written,

$$\langle n_v \rangle = \frac{1}{Q_{GC}} \sum_{n_v} \sum_{i} n_v g_i \, e^{-\beta E_{i,n_v} + \beta n_v \mu}, \tag{9}$$

where the grand canonical partition function is defined by

$$Q_{GC} = \sum_{n_v} \sum_{i} g_i \, e^{-\beta E_{i,n_v} + \beta n_v \mu}. \tag{10}$$

E_{i,n_v} and g_i are the energies and degeneracies of states we denote as inside the vacancy respectively, $\mu(T, y)$ is the chemical potential of hydrogen, to be calculated in the bulk reservoir, in which y denotes the bulk concentration, in units to be discussed below. We emphasize that 'inside the vacancy' means inside a rather small supercell containing the vacancy, the exact size of which is arbitrary. β is $1/(k_{\mathrm{B}}T)$, where k_{B} and T are the Boltzmann constant and temperature. Eq (9) can be evaluated over the regimes of interest by writing it in terms of the canonical partition function

$$Q_C^{vac}(T, n_v) = \sum_{i} g_i e^{-\beta E_{i,n_v}}, \tag{11}$$

which we can substitute into (9) to give

$$\langle n_v \rangle = \frac{1}{Q_{GC}} \sum_{n_v} n_v Q_C^{vac}(T, n_v) \, e^{\beta n_v \mu}$$

in which

$$Q_{GC} = \sum_{n_v} Q_C^{vac}(T, n_v) \, e^{\beta n_v \mu}. \tag{12}$$

This manipulation conveniently splits our task into two distinct parts. Firstly we need to calculate $\mu(T, y)$ for any low concentration y of bulk H, and secondly we need calculations of $Q_C^{vac}(T, n_v)$ for integer values of n_v from 0 to 6 (more than 6 will not be relevant). We can then combine the two kinds of calculation into formula (12). The calculation of Q_C^{vac} is the task of our nested sampling, so it remains to derive an expression for μ.

Assuming zero pressure, we shall evaluate the Helmholtz free energy of a dilute concentration y hydrogen atoms per Fe atom. Let us divide a perfect bulk crystal into N identical space-filling cells, each containing N_a atoms of Fe. Normally we have in mind here the supercells with periodic boundary conditions, as used in DFT calculations, in which $1 \leq N_a \leq 54$. We can apply the following approach both to a lattice model, in which only tetrahedral sites are occupied, or the continuous phase space which is used in this paper. The thermodynamic limit of $N \to \infty$ will be taken.

We suppose the bulk hydrogen to be distributed over these cells such that there is at most one H atom per cell, which is an excellent approximation for our purposes as indicated above. The number of Fe atoms is $N_a N$ and the number of H atoms is $n = yN_aN$, which is equal to the number of supercells occupied by a hydrogen atom. The total partition function Z_C for the perfectly crystalline bulk system (remember we are neglecting here the movement of the Fe lattice) can therefore be factorised into the partition function of n singly-occupied cells and the partition function for the choice of these n cells from the total N:

$$Z_C = Q_C^{bulk}(T, 1/N_a)^n \cdot \frac{N!}{(N-n)!n!}\,. \tag{13}$$

The chemical potential will be obtained from this expression via the standard thermodynamic formula:

$$\mu = -k_{\mathrm{B}}T\frac{\partial}{\partial n}\ln Z_C. \tag{14}$$

We make the usual Stirling's approximation for the factorials in Eqn (13), namely

$$\ln\left(\frac{N!}{(N-n)!n!}\right) \to NS(n/N) \tag{15}$$

where

$$S(x) = -x\ln(x) - (1-x)\ln(1-x), \tag{16}$$

which is accurate in our thermodynamic limit. Inserting this into (14), the derivative is straightforward and we obtain

$$\mu = k_{\mathrm{B}}T\ln\frac{n/N}{1-n/N} - k_{\mathrm{B}}T\ln Q_C^{bulk}(T, 1/N_a). \tag{17}$$

Replacing n/N by yN_a, this can be written in terms of the concentration per host atom in the form

$$\mu = k_{\mathrm{B}}T\ln(yN_a) - k_{\mathrm{B}}T\ln Q_C^{bulk}(T, 1/N_a), \tag{18}$$

where we have assumed again $yN_a \ll 1$ in order to make the replacement $1 - N_a y \to 1$ in the first term. Consider now the calculation of the second term using a supercell and periodic boundary conditions with nested sampling or any other sampling technique. As long as N_a is big enough to avoid spurious interactions between the single H atom and its periodic images, it is clear that

$$Q_C^{bulk}(T, 1/N_a) = N_a Q_C^{bulk}(T, 1). \tag{19}$$

This linear scaling ensures that N_a cancels out of expression (18), which we can finally write as

$$\mu = k_{\mathrm{B}} T \ln(y) - k_{\mathrm{B}} T \ln Q_C^{bulk}(T, 1), \tag{20}$$

from which

$$\exp(\beta\mu) = y / Q_C^{bulk}(T, 1). \tag{21}$$

The grand canonical partition function for hydrogen in the vacancy can now be expressed as

$$Q_{GC} = \sum_{n_v} Q_C^{vac}(T, n_v) \left(y / Q_C^{bulk}(T, 1)\right)^{n_v} \tag{22}$$

and the mean number of trapped atoms is

$$\langle n_v \rangle = \frac{1}{Q_{GC}} \sum_{n_v} n_v \, Q_C^{vac}(T, n_v) \left(y / Q_C^{bulk}(T, 1)\right)^{n_v}. \tag{23}$$

Both canonical partition functions are computed by nested sampling, which we apply to the full energy landscape and, for comparison purposes only, to the lattice model, both with and without H-H interactions. Taking the reference state for hydrogen to be when it is at rest in a bulk tetrahedral site, Eq. (12) still applies to the discrete lattice models, with the simplification that $Q_C^{bulk}(T, 1)$ reduces to 6. For the 6 vacancy trapping sites, in our lattice model calculations we take the 9 non-degenerate energies for the states of $1 \leq n_v \leq 6$ from our static DFT calculations. The effect of the H-H interactions is to split the degeneracy of all the states with $2 \leq n_v \leq 5$. While the single vacancy with a few H atoms comprise a relatively simple system in terms of composition, size and geometry, the same approach could be applied to any system in which a solute has a low concentration in the bulk, so

that the defective and perfectly crystalline regions can be considered separately using the nested sampling method and linked via the explicit calculation of the partition functions and the chemical potential.

V. RESULTS

The average number of H atoms in a vacancy is shown in Fig. 3(a) as a function of temperature for a range of hydrogen concentrations, calculated using nested sampling, the grand canonical approach, and the GAP model for the potential energy surface. For comparison we show the same for the lattice model in Fig. 3(b). The expected occupancy of the vacancy using Oriani's theory, which does not include H-H interactions, is shown in Fig. 3(c). From these results we can estimate the temperature at which the vacancy goes from being fully to half occupied (6H to 3H). Our model (Fig. 3(a)) indicates this happens at 340 K for a bulk H concentration of 0.01 ppm, or 500 K for the very high concentration of 10 ppm. The reduced phase space of the lattice model (Fig. 3(b)) spuriously enhances the trapping, so that the corresponding temperatures are 380 and 630 K. By ignoring H-H repulsion we further increases the apparent occupancy of the vacancy, bringing the temperature of half-occupancy up to 450 and 750 K for 0.01 ppm and 10 ppm respectively (Fig. 3(c)). As noted before, we kept the positions of Fe atoms fixed in the DFT optimal configuration. Although certainly possible within the same methodological framework, allowing movement of Fe atoms is not expected to change our conclusions significantly, since their vibrations are nearly harmonic at room temperature and their contribution to configurational entropy is much smaller than that of the mobile H atoms.

VI. CONCLUSIONS

Using a purpose-made hydrogen-iron interaction potential defined by Gaussian process regression fitted to DFT calculations, we applied nested sampling to obtain the mean number of H-atoms trapped in a vacancy in an α-Fe crystal, as a function of bulk hydrogen concentrations from 10^{-4} down to 10^{-11}. To deal with such low bulk hydrogen concentrations we introduce a grand-canonical formulation, which requires a separate calculation of the

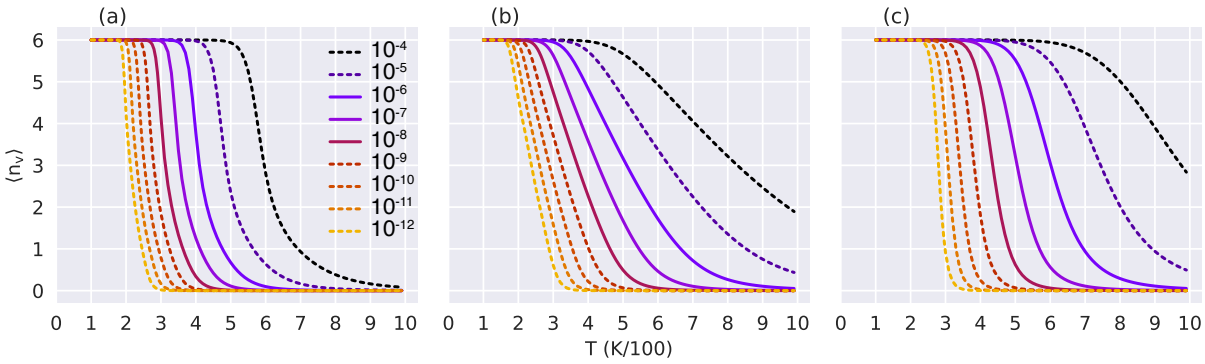

FIG. 3. Expected number of H-atoms in the vacancy as a function of temperature for a range of bulk H concentrations, calculated using (a) the main model presented in this work, (b) the same nested sampling approach applied to the discrete lattice system, in which the energies are only required for H atoms on the bulk tetrahedral sites and the six sites in the vacancy, and (c) the Oriani model, using the same single H on-site energies as (b), but switching off the H-H interactions. The concentrations are in units of atom fraction. The three solid lines represent concentrations most relevant to industrially observed H embrittlement of steels.

chemical potential of H as a function of its bulk concentration. We obtained this chemical potential from nested sampling calculations on a single H atom in the perfect crystal. Our grand canonical approach to dilute solutions does not suffer from the difficulty of simulation methods for which the required supercell volume scales inversely with the concentration of solute.

At sufficiently low temperature, the saturated vacancy trap for all bulk concentrations contains 6 H. Our calculations indicate that this maximum level of trapping would *not* be achieved at around room temperature for bulk concentrations below 0.1 atom ppm. It is widely accepted that the propensity for steels to embrittle is correlated with the amount of "mobile hydrogen" rather than the total including those in deep traps such as carbides, grain boundaries etc. and thus understanding the nature and dynamics of mobile hydrogen is in turn critical for understanding hydrogen embrittlement. The designation of "mobile hydrogen" is conventionally taken to be equivalent to H atoms in bulk interstitial lattice sites, and sometimes referred to as "lattice hydrogen". Our results however suggest that the H atoms trapped at vacancies contribute to the pool of mobile hydrogen in the system. Thus theories of hydrogen dynamics in iron and steel, including analyses of thermal desorption

spectroscopy (the main analytical tool in the field) such as Kissinger theory[54] likely have to be revised. There will be further direct implications to theories of H-embrittlement that involve the interaction with vacancies explicitly, such as "hydrogen enhanced strain induced vacancy formation"[1].

For the present work we assumed classical nuclei, which would be a poor approximation for hydrogen in metals much below room temperature. The approach could readily be extended to include nuclear quantum effects for H using the Feynman path-integral formalism which replaces the H-atoms with ring-polymer chains. See for example Gillan *et al.* [55] for a very clear exposition. We similarly ignored movement of Fe atoms, and diffusion leading to clustering of the Fe monovacancies. The latter would lead to a variety of other trapping configurations, from divacancies upwards, which we have not discussed here. The same treatment could be applied to all these aspects of the problem, using the GAP model for Fe-Fe of Dragoni *et al.* [48]. Our combination of machine-learned interatomic potentials and nested sampling, together with our grand canonical approach to trap occupancy, should be useful for many similar questions involving segregation and trapping of impurities at interfaces or dislocations.

We acknowledge the support of the UK EPSRC under the Programme Grant HEmS, EP/ L014742, and an anonymous referee who helped us to greatly improve an earlier version of this paper.

---

[1] M. Nagumo, Hydrogen related failure of steels - a new aspect, Mater Sci Technology **20**, 940 (2004).

[2] Y. Fukai, *The metal-hydrogen system: basic bulk properties*, Vol. 21 (Springer Science & Business Media, 2006).

[3] T. Neeraj, R. Srinivasan, and J. Li, Hydrogen embrittlement of ferritic steels: Observations on deformation microstructure, nanoscale dimples and failure by nanovoiding, Acta Materialia **60**, 5160 (2012).

[4] R. Gangloff and B. Somerday, *Gaseous hydrogen embrittlement of materials in energy technologies: the problem, its characterisation and effects on particular alloy classes* (Elsevier, 2012).

[5] A. T. Paxton, A. P. Sutton, and M. W. Finnis, eds., *The Challenges of Hydrogen and Metals*, Philosophical Transactions A: Mathematical, Physical and Engineering Sciences, Vol. 375 (The Royal Society, London, 2017).

[6] J. Takahashi, K. Kawakami, and T. Tarui, Direct observation of hydrogen-trapping sites in vanadium carbide precipitation steel by atom probe tomography, Scripta Materialia **67**, 213 (2012).

[7] Y.-S. Chen, D. Haley, S. S. Gerstl, A. J. London, F. Sweeney, R. A. Wepf, W. M. Rainforth, P. A. Bagot, and M. P. Moody, Direct observation of individual hydrogen atoms at trapping sites in a ferritic steel, Science **355**, 1196 (2017).

[8] E. Hayward and C.-C. Fu, Interplay between hydrogen and vacancies in alpha-Fe, Physical Review B **87**, 174103 (2013).

[9] D. Mirzaev, A. Mirzoev, K. Y. Okishev, A. Shaburov, G. Ruzanova, and A. Ursaeva, Formation of hydrogen-vacancy complexes in alpha iron, Phys. Met. Metall. **113**, 923 (2012).

[10] K. Ohsawa, K. Eguchi, H. Watanabe, M. Yamaguchi, and M. Yagi, Configuration and binding energy of multiple hydrogen atoms trapped in monovacancy in bcc transition metals, Physical Review B **85**, 094102 (2012).

[11] H. Sugimoto and Y. Fukai, Hydrogen-induced superabundant vacancy formation in bcc fe: Monte carlo simulation, Acta Materialia **67**, 418 (2014).

[12] Y. Tateyama and T. Ohno, Stability and clusterization of hydrogen-vacancy complexes in alpha-Fe: An ab initio study, Physical Review B **67**, 174105 (2003).

[13] Fukai, Yuk and Ōkuma, Nobuyuki, Evidence of Copious Vacancy Formation in Ni and Pd under a High Hydrogen Pressure, Japanese Journal of Applied Physics **32**, L1256 (1993).

[14] D. Tanguy and M. Mareschal, Superabundant vacancies in a metal-hydrogen system: Monte carlo simulations, Physical Review B **72**, 174116 (2005).

[15] R. Kirchheim, Reducing grain boundary, dislocation line and vacancy formation energies by solute segregation. i. theoretical background, Acta Materialia **55**, 5129 (2007).

[16] R. Kirchheim, On the solute-defect interaction in the framework of a defactant concept, International Journal of Materials Research **100**, 483 (2009).

[17] L. Ismer, M. S. Park, A. Janotti, and C. G. Van de Walle, Interactions between hydrogen impurities and vacancies in Mg and Al: A comparative analysis based on density functional theory, PHYSICAL REVIEW B **80**, 10.1103/PhysRevB.80.184110 (2009).

[18] R. Nazarov, T. Hickel, and J. Neugebauer, First-principles study of the thermodynamics of hydrogen-vacancy interaction in fcc iron, Physical Review B **82**, 224104 (2010).

[19] R. Nazarov, T. Hickel, and J. Neugebauer, Ab initio study of H-vacancy interactions in fcc metals: Implications for the formation of superabundant vacancies, PHYSICAL REVIEW B **89**, 10.1103/PhysRevB.89.144108 (2014).

[20] L. Bukonte, T. Ahlgren, and K. Heinola, Thermodynamics of impurity-enhanced vacancy formation in metals, JOURNAL OF APPLIED PHYSICS **121**, 10.1063/1.4974530 (2017).

[21] A. Metsue, A. Oudriss, and X. Feaugas, Thermodynamic vs. Kinetic Origin of Superabundant Vacancy Formation in Ni Single Crystals, Corrosion **75**, 898 (2019).

[22] J. Hou, X.-S. Kong, X. Wu, J. Song, and C. Liu, Predictive model of hydrogen trapping and bubbling in nanovoids in bcc metals, Nature Materials **18**, 833 (2019).

[23] R. Kirchheim, Revisiting hydrogen embrittlement models and hydrogen-induced homogeneous nucleation of dislocations, Scripta materialia **62**, 67 (2010).

[24] J. D. Silva and R. B. Mclellan, The solubility of hydrogen in super-pure-iron single crystals, J. Less-Common Met. **50**, 1 (1976).

[25] R. A. Oriani, The diffusion and trapping of hydrogen in steel, Acta Metall. Mater. **18**, 147 (1970).

[26] D. Tanguy, Y. Wang, and D. Connetable, Stability of vacancy-hydrogen clusters in nickel from first-principles calculations, Acta Materialia **78**, 135 (2014).

[27] D. Jiang and E. A. Carter, Diffusion of interstitial hydrogen into and through bcc fe from first principles, Phys. Rev. B **70**, 064102 (2004).

[28] D. C. Sorescu, First principles calculations of the adsorption and diffusion of hydrogen on fe (100) surface and in the bulk, Catalysis today **105**, 44 (2005).

[29] P. Hohenberg and W. Kohn, Inhomogeneous electron gas, Phys. Rev **136**, B864 (1964).

[30] W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, Phys. Rev **140**, A1133 (1965).

[31] J. Behler and M. Parrinello, Generalized neural-network representation of high-dimensional potential-energy surfaces, Phys. Rev. Lett. **98**, 146401 (2007).

[32] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons, Phys. Rev. Lett. **104**, 136403 (2010).

[33] A. P. Thompson, L. P. Swiler, C. R. Trott, S. M. Foiles, and G. J. Tucker, Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials, Journal of Computational Physics **285**, 316 (2015).

[34] A. V. Shapeev, Moment Tensor Potentials: A Class of Systematically Improvable Interatomic Potentials, Multiscale Modeling & Simulation **14**, 1153 (2016).

[35] C. Micheletti, A. Laio, and M. Parrinello, Reconstructing the density of states by history-dependent metadynamics, Phys. Rev. Lett. **92**, 170601 (2004).

[36] F. Wang and D. Landau, Efficient, multiple-range random walk algorithm to calculate the density of states, Phys. Rev. Lett. **86**, 2050 (2001).

[37] R. H. Swendsen and J.-S. Wang, Replica monte carlo simulation of spin-glasses, Phys. Rev. Lett. **57**, 2607 (1986).

[38] D. Frantz, D. L. Freeman, and J. Doll, Reducing quasi-ergodic behavior in monte carlo simulations by j-walking: Applications to atomic clusters, J. Chem. Phys. **93**, 2769 (1990).

[39] J. Skilling, Nested sampling, AIP Conference Proceedings **735**, 395 (2004).

[40] J. Skilling, Nested sampling for general Bayesian computation, Bayesian analysis **1**, 833 (2006).

[41] S. Fujikake, V. L. Deringer, T. H. Lee, M. Krynski, S. R. Elliott, and G. Csányi, Gaussian approximation potential modeling of lithium intercalation in carbon nanostructures, J. Chem. Phys. **148**, 241714 (2018).

[42] S. T. John and G. Csanyi, Many-body coarse-grained interactions using gaussian approximation potentials, J. Phys. Chem. B **121**, 10934 (2017).

[43] A. P. Bartók and G. Csányi, G aussian approximation potentials: A brief tutorial introduction, Int. J. Quant. Chem. **115**, 1051 (2015).

[44] P. Rowe, G. Csányi, D. Alfè, and A. Michaelides, Development of a machine learning potential for graphene, Phys. Rev. B **97**, 054303 (2018).

[45] V. L. Deringer, C. J. Pickard, and G. Csányi, Data-driven learning of total and local energies in elemental boron, Phys. Rev. Lett. **120**, 156001 (2018).

[46] V. L. Deringer, C. Merlet, Y. Hu, T. H. Lee, J. A. Kattirtzi, O. Pecher, G. Csányi, S. R. Elliott, and C. P. Grey, Towards an atomistic understanding of disordered carbon electrode materials, Chem. Comm. (2018).

[47] M. Ceriotti, M. J. Willatt, and G. Csányi, Machine Learning of Atomic-Scale Properties Based on Physical Principles, in *Handbook of Materials Modeling: Methods: Theory and Modelling*,

edited by A. W. and Y. S. (Springer, Cham, 2018) pp. 1–27.

[48] D. Dragoni, T. D. Daff, G. Csányi, and N. Marzari, Achieving dft accuracy with a machine-learning interatomic potential: Thermomechanics and defects in bcc ferromagnetic iron, Phys. Rev. Materials **2**, 013808 (2018).

[49] A. Ramasubramaniam, M. Itakura, and E. A. Carter, Interatomic potentials for hydrogen in α–iron based on density functional theory, Phys. Rev. B **79**, 174101 (2009).

[50] *QUIP* (2017, Last checked 08 May 2020), `http://github.com/libatoms/quip`.

[51] See Supplemental Material at [URL will be inserted by publisher] for more details and values of the parameters for these potentials.

[52] L. B. Pártay, A. P. Bartók, and G. Csányi, Efficient sampling of atomic configurational spaces, J. Phys. Chem. B **114**, 10502 (2010).

[53] *PYMATNEST* (2016, Last checked 08 May 2020), `http://github.com/libatoms/pymatnest`.

[54] H. E. Kissinger, Variation of peak temperature with heating rate in differential thermal analysis, Journal of Research of the National Bureau of Standards **57**, 217 (1956).

[55] M. J. Gillan, The path-integral simulation of quantum systems, in *Computer Modelling of Fluids Polymers and Solids*, Vol. 293, edited by C. R. A. Catlow, C. S. Parker, and M. P. Allen (Kluwer Academic Publ, 1990) pp. 155–188.

APPENDIX: TEST OF THE 1 H/SUPERCELL APPROXIMATION

In order to appreciate better our approximation of no more than 1 H atom per perfect supercell we can demonstrate it with the lattice model, assuming an ideal solution of H on a lattice of tetrahedral sites, of which there are six per Fe atom. In particular, we examine the error made by our approximation as a fraction of the exact entropy in the ideal solution. We make use of the function $S(x)$, defined in Eq (15), which is the familiar dilute solution expression for the entropy per site when the fractional site-occupancy is x. We also work with the number of atoms per supercell, N_a, and the number of sites per atom, N_s, which in this example are to be 54 and 6 respectively, and the number of H atoms per host atom is $y \ll 1$, as above. The 'exact' dilute solution entropy per atom is then

$$S_{\text{exact}} = N_s S(y/N_s) \tag{24}$$

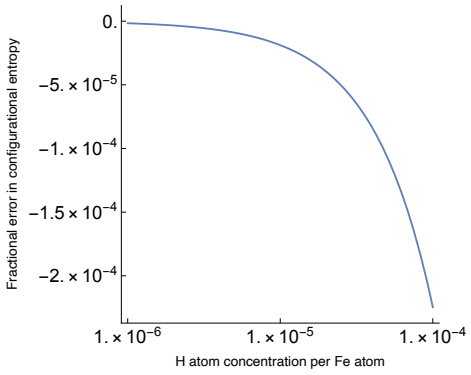

FIG. 4. Estimate of the fractional error in the approximated configurational entropy of bulk hydrogen with our 1H-per-supercell model. This error estimate is based on the lattice approximation, in which H can only occupy tetrahedral sites.

In the lattice model, our approximate partition function Eq (13) becomes

$$Z_C^{\text{lattice}} = (N_s N_a)^n \cdot \frac{N!}{(N-n)!n!}\,. \tag{25}$$

From which the entropy per atom is

$$\begin{aligned} S_{\text{approx}} &= \frac{1}{NN_a} \ln Z_C^{\text{lattice}} \\ &= \frac{n}{NN_a} \ln(N_s N_a) + \frac{1}{N_a} S(n/N) \\ &= y \ln(N_s N_a) + \frac{1}{N_a} S(yN_a). \end{aligned} \tag{26}$$

We have made use of the identity $yN_a = n/N$ to eliminate the extensive quantities N and n. The error is evaluated as $S_{\text{approx}} - S_{\text{exact}}$. It is easy to see by Taylor expanding in y, that the terms to leading order as $y \to 0$, namely $y \ln y$, cancel, and the fractional error is very small until y exceeds 10^{-4}. This is shown graphically in Fig. 4.