

## ARTICLE OPEN



## Atomistic fracture in bcc iron revealed by active learning of Gaussian approximation potential

Lei Zhang<sup>1</sup>✉, Gábor Csányi<sup>2</sup>, Erik van der Giessen<sup>3</sup> and Francesco Maresca<sup>1</sup>✉

The prediction of atomistic fracture mechanisms in body-centred cubic (bcc) iron is essential for understanding its semi-brittle nature. Existing atomistic simulations of the crack-tip under mode-I loading based on empirical interatomic potentials yield contradicting predictions and artificial mechanisms. To enable fracture prediction with quantum accuracy, we develop a Gaussian approximation potential (GAP) using an active learning strategy by extending a density functional theory (DFT) database of ferromagnetic bcc iron. We apply the active learning algorithm and obtain a Fe GAP model with a converged model uncertainty over a broad range of stress intensity factors (SIFs) and for four crack systems. The learning efficiency of the approach is analysed, and the predicted critical SIFs are compared with Griffith and Rice theories. The simulations reveal that cleavage along the original crack plane is the atomistic fracture mechanism for {100} and {110} crack planes at  $T = 0$  K, thus settling a long-standing issue. Our work also highlights the need for a multiscale approach to predicting fracture and intrinsic ductility, whereby finite temperature, finite loading rate effects and pre-existing defects (e.g., nanovoids, dislocations) should be taken explicitly into account.

npj Computational Materials (2023)9:217; <https://doi.org/10.1038/s41524-023-01174-6>

## INTRODUCTION

Brittle fracture is a key failure mechanism of body-centred cubic (bcc) transition metals, which limits their application and can jeopardise the safety of infrastructures. Brittle fracture usually takes place in the form of cleavage, which can be accompanied by dislocation plasticity. The competition between thermally activated dislocation mobility and (atomic-scale) crack tip deformation mechanisms controls the fracture process of bcc iron<sup>1–4</sup>. Experiments on single-crystal bcc iron reveal that cleavage takes place on {100} planes for pre-existing {100} and {110} crack planes with ⟨100⟩ and ⟨110⟩ crack fronts<sup>5</sup>. However, atomic-scale experimental data is unavailable for crack tips in bcc iron, leaving the atomistic crack-tip mechanisms unclear.

Atomistic modelling based on molecular dynamics (MD) approach has been widely used to predict the atomistic fracture mechanisms<sup>6–15</sup> (see *Supplementary: A brief summary of existing atomistic studies on fracture in bcc iron*). Mode-I atomistic crack tip simulations at  $T = 0$  K are typically used to assess the intrinsic ductility of metals, which is controlled by the competition between crack propagation (cleavage) and dislocation emission from the crack tip<sup>3,4,16</sup>. Yet, as highlighted in extensive reviews<sup>14,15</sup>, classical interatomic potentials (IAPs) for bcc iron predict contradicting crack-tip mechanisms at  $T = 0$  K (i.e., cleavage, crack propagation planes, dislocation emission, and phase transition) for the same crack system. Despite not being explicitly designed for fracture, some of these potentials have been widely used to simulate cracks<sup>6–15</sup>. Yet, none of the empirical IAPs agree with the low-temperature experimental fracture mechanisms as a function of the pre-existing crack system<sup>5</sup>.

Figure 1 shows the critical stress intensity factor ( $K_{lc}$ ) under mode-I loading as computed by performing molecular statics (MS) simulations using seven classical IAPs (see *MS/MD simulation setup* in *Methods*) that were not investigated in ref. <sup>14</sup> nor ref. <sup>15</sup>, including two embedded atom method (EAM) potentials<sup>17,18</sup>,

three Modified EAM (MEAM) potentials<sup>19–21</sup>, one bond order potential<sup>22</sup>, and one angular dependent potential<sup>23</sup>. Results are compared with the critical stress intensity factors according to Griffith theory<sup>24</sup>

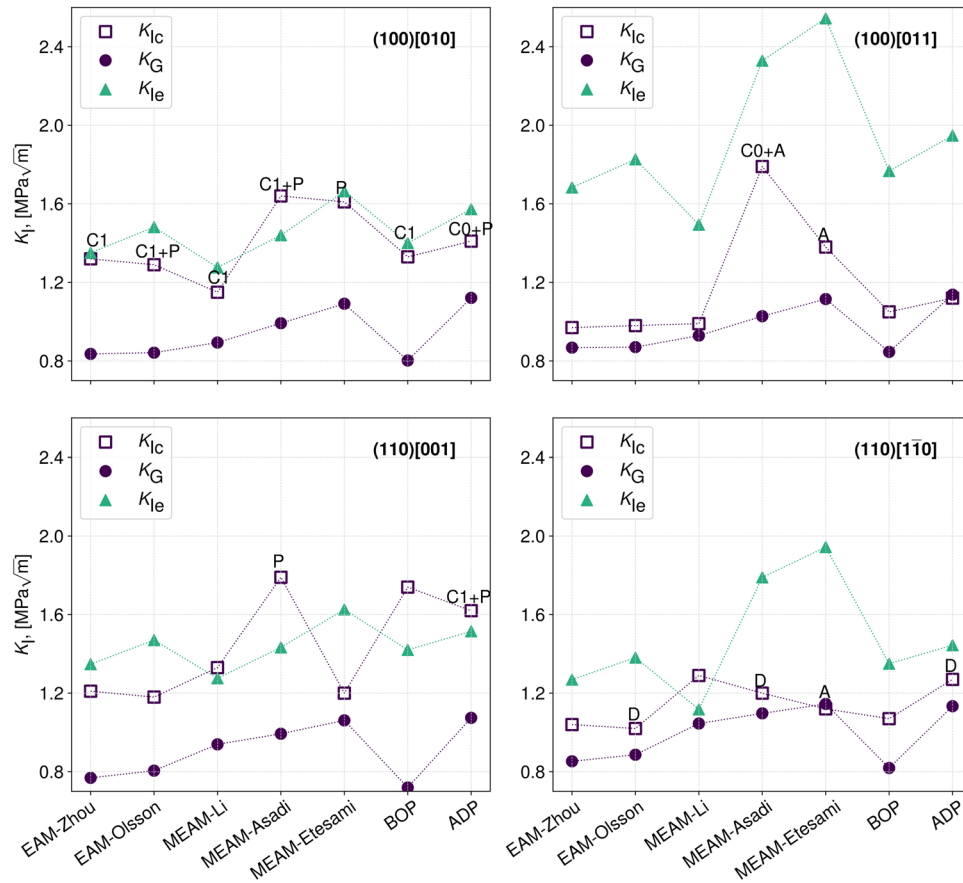
$$K_G = \sqrt{\frac{2\gamma_s}{B}}, \quad (1)$$

where  $\gamma_s$  is the surface energy and  $B$  is a constant determined by the elastic constants (see *Supplementary Eqs. 9–17*). Dislocation emission is predicted according to Rice theory<sup>2</sup> with an anisotropic implementation<sup>25</sup>

$$K_{le} = \frac{\sqrt{\gamma_{usf} o(\theta, \phi)}}{F_{12}(\theta)}, \quad (2)$$

where  $o(\theta, \phi)$  is a function of  $\theta$  and  $\phi$ .  $\theta$  is the angle between slip plane and crack plane,  $\phi$  is the angle between the slip direction and a vector lying on the slip plane and perpendicular to the crack-front direction, and  $\gamma_{usf}$  is the unstable stacking fault energy of the slip plane. The computational details of  $o(\theta, \phi)$  and  $F_{12}(\theta)$  are given in *Supplementary Eqs. 18–30*. These potential-dependent properties (elastic constants,  $\gamma_s$  and  $\gamma_{usf}$ ) used for computing  $K_G$  and  $K_{le}$  are listed in *Supplementary Table 2*. For (100)[010] (crack plane/crack front) crack system, no IAP predicts pure cleavage on the original crack plane, in stark contrast with both the low-temperature experimental observations<sup>5</sup> and the theoretical predictions ( $K_G$  and  $K_{le}$ ). Instead, all IAPs predict either cleavage on {110} planes (rather than {100} planes), or phase transition, or cleavage accompanied by phase transition. With only two exceptions (MEAM-Asadi and MEAM-Etesami), the predictions of (100)[011] are in good agreement with experiments<sup>5</sup> and Griffith theory. For (110)[001], no IAP is able to predict pure cleavage on {100} plane. Cleavage and dislocation emission are both observed in (110)[1 $\bar{1}$ 0] crack system. In summary, most classical IAPs predict cleavage for (100)[011] and (110)[001] crack

<sup>1</sup>Engineering and Technology Institute Groningen, Faculty of Science and Engineering, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands. <sup>2</sup>Engineering Laboratory, University of Cambridge, Cambridge CB2 1PZ, UK. <sup>3</sup>Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands. ✉email: lei.zhang@rug.nl; f.maresca@rug.nl



**Fig. 1** Summary of fracture simulations for crack system (100)[010], (100)[011], (110)[001], and (110)[110].  $K_{Ic}$  is the critical stress intensity factor predicted by MS simulations. The letters above the symbols indicate the fracture mechanism observed in MS simulations with the different IAPs. **C0** and **C1** indicate cleavage on {100} and {110} planes that are different from the original crack plane. **P** and **D** indicate phase transformation and dislocation emission from the crack tip, respectively. **A** indicates amorphous structure forming at the crack tip. The crack propagates on the original crack plane if no symbol is specified.  $K_G$  and  $K_{Ie}$  are predictions according to Griffith<sup>24</sup> and Rice theories<sup>2</sup> for cleavage and dislocation emission, respectively.

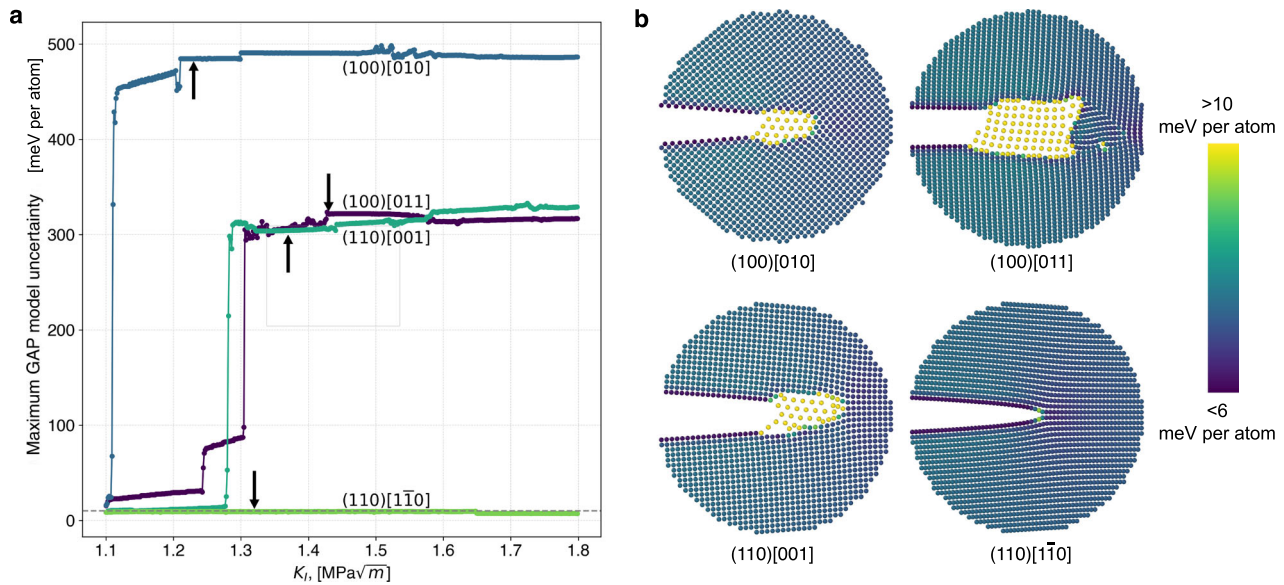
systems while yielding conflicting results for the two remaining crack systems, thus leaving the question of what atomistic mechanisms control crack-tip behaviour of bcc iron unresolved.

The lack of agreement between atomistic predictions and experiments of fracture modes and critical stress intensity factors motivates the development of a potential that is capable of reproducing the experimental observations. Machine learning (ML) potentials enable MD/MS simulation to describe a complex potential energy surface (PES) with quantum accuracy, and are orders of magnitude faster than density functional theory (DFT). Among the existing ML potential frameworks, Gaussian Approximation Potential (GAP) has been shown to accurately describe the complex motion of screw dislocations in bcc iron<sup>26</sup> and the cleavage process in silicon<sup>27</sup>. As a Gaussian process regression method, GAP can predict both the mean value and the variance of the atomic energies<sup>28</sup>. GAP predicted variance can be used as an indicator for extrapolation, which is practical to evaluate the model uncertainty in iterative/active learning process<sup>29,30</sup>.

We first consider an existing GAP for ferromagnetic bcc iron<sup>31</sup>, hereafter named Fe-GAP18. The Fe-GAP18 database has been developed to study thermodynamics and defects in bcc iron<sup>32</sup>. The database includes stretched primitive cells, point defects, interfaces etc. Using the GAP model uncertainty, we show (Fig. 2a) that Fe-GAP18 cannot predict the fracture behaviour of bcc iron with a converged uncertainty. In particular, Fig. 2a reports the maximum model uncertainty of crack-tip atoms during mode-I MS

simulations at multiple applied  $K_I$ 's. Only crack propagation along (110)[110] is predicted with a converged model uncertainty. Figure 2b shows snapshots with typical artifacts produced during the fracture process. We argue that Fe-GAP18 is unable to capture the fracture behaviour due to the lack of crack tip atomic environments in the DFT training dataset rather than due to an intrinsic limitation of GAP. Therefore, we develop a systematic strategy to improve GAP for fracture predictions, including a preliminary fracture-relevant database extension, which is subsequently supplemented with an active-learned database. We show that the extended database enables GAP to predict atomistic fracture mechanisms with a converged model uncertainty at  $T=0K$  and at finite temperatures.

The paper is organised as follows. We describe the approach and the active learning algorithm used to expand the DFT database in section *Results "Preparation of the Reference Database"*. In section *Results "Implementation of active learning"*, we implement the active learning algorithm and show the convergence of our approach (hereafter indicated as Fe-GAP22). Based on Fe-GAP22, we predict the fracture behaviour of single crystal bcc iron at  $T=0K$  and at finite temperatures under high loading rate conditions in section *Results "Prediction of fracture mechanisms and critical stress intensity factors ( $K_{Ic}$ )"*. In section *Discussion*, we compare the model with previous reviews and experiments. We conclude by outlining further research directions to be explored with the quantum-accurate Fe-GAP22.



**Fig. 2** Fracture predictions of Fe-GAP18 trained on Dragoni et al. original database<sup>31</sup>. **a** Maximum model uncertainty during fracture simulations of four crack systems. The dashed line indicates the 10 meV per atom, and the arrows indicate the configuration shown in **(b)**. **b** Simulation snapshots coloured by the model uncertainty. Note that this is a new version of Fe-GAP18 trained on the database by Dragoni et al. using an optimised descriptor<sup>37</sup> (see section *Methods: GAP training*).

## RESULTS

### Preparation of the reference database

We first extended the existing DFT database of Fe-GAP18<sup>32</sup> with primitive cells strained according to the large deformations that may occur at crack tips. The strain states near a crack tip under plane strain and plane stress conditions are estimated from anisotropic classical linear elastic (CLE) crack-tip fields. The maximum tensile strain is found to be 0.27 at the point that is 0.1 nm away from the crack-tip under  $K_I = 1.5 \text{ MPa}\sqrt{\text{m}}$ . By analysing the strain components of the primitive cell from the original database<sup>32</sup> comprising 6000 distorted primitive bcc cells, we found that 95% of original data is within the strain of 0.15 and the rest are highly stretched along one crystal direction (2–4 times). We therefore created an additional set of highly strained primitive cells and enriched the DFT dataset to encompass the anisotropic CLE predicted strain states, referred to as DB9. Note that the original database from Dragoni et al.<sup>32</sup> includes 8 subsets, referred to as DB1–DB8. In order to ensure an unbiased sampling of atomic environments, we employed random uniform sampling algorithm *SOBOL*<sup>33</sup> instead of the strain states directly obtained by anisotropic CLE crack-tip fields (see *Supplementary: SOBOL sampling approach*). Principal component analysis of the strain components shows that DB9 expands the original database and encompasses a more extensive set of strain states than those calculated by anisotropic CLE theory (see Supplementary Fig. 1). Common neighbour analysis shows that the original database and DB9 also include face-centred cubic (fcc) structures. Besides, we included distorted hexagonal close-packed (hcp) primitive cells (referred to as DB10) to ensure that GAP is capable of predicting the hcp phase.

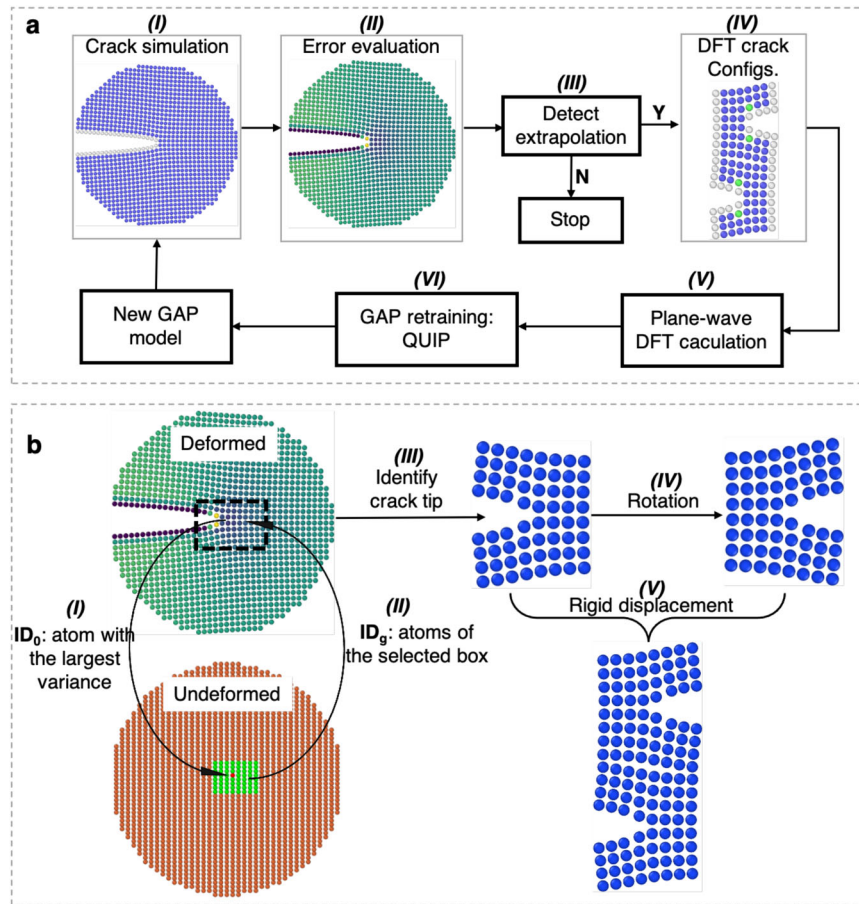
The magnetic contribution to the total energy in iron plays an important role and thus cannot be neglected. Our DFT calculation of the primitive bcc cell shows that the magnetisation degree decreases when the volume decreases. The primitive cell becomes completely non-magnetic when the volume reduction is larger than 37.5%. Including those magnetic states beyond ferromagnetic deteriorates the performance of GAP, leading to unrealistic predictions, such as negative surface energy. This is because the transition between ferromagnetic and other magnetic states (e.g., non-magnetic) is not unique (i.e., not smooth and not single-

valued), which gives rise to a partly spurious GAP predicted PES since magnetism is not explicitly accounted for in the current GAP implementation. Depending on the magnetisation parameter settings in DFT calculations, the self-consistent calculation may be stuck in a ground state of ferromagnetism that has larger energy than the non-magnetic ground state, or vice-versa. The mixture of configurations with the same atomic environment and distinct magnetic states leads to a spurious GAP fitted PES. Here, we make the assumption that the crack tip stays ferromagnetic, since the locally highly strained crack-tip bonds are surrounded by ferromagnetic iron atoms; for this reason, we train Fe-GAP22 on a consistent database of ferromagnetic configurations. Furthermore, DFT calculations used within the active learning scheme (see next Section) show that the crack-tip atoms stay ferromagnetic (see Supplementary Fig. 2).

Bond rupture and dislocation emission from the crack tip are two fundamental crack-tip mechanisms<sup>34</sup>. The DFT database should be sufficiently informed with configurations relevant to those mechanisms. Maresca et al. showed that Fe-GAP18 predicts the single-hump Peierls potential and the compact core structure of screw dislocations with DFT accuracy<sup>26</sup>. However, Fe-GAP18 does not include the separation process of atomic planes explicitly. Here, we added rigid separation configurations of {100}, {110}, {111} and {112} crystallographic planes for thin and thick slabs (referred to as DB11). The rigid separation includes two sets, i.e., ideal crystal structures and “rattled” structures. The rattled structures are obtained by adding a Gaussian noise  $\mathcal{N} \sim [0, 0.05a_0]$  ( $a_0$  lattice constant) to the three Cartesian coordinates of all atoms to explore the PES beyond highly symmetric, minimum energy configurations. In either case, separations are performed from 0 to 4 Å with a step of 0.4 Å.

The detailed information about the preliminary database is listed in Supplementary Table 3. As shown in Fig. 4, from the original database (“Original”) to preliminary database (“Iter-0”), adding those relevant configurations considerably reduces the maximum model uncertainty during the fracture simulations. Yet, GAP based on “Iter-0” database still cannot predict the fracture process with a converged model uncertainty (see Supplementary Fig. 3).

Since the crack tip deformation fields are inhomogeneous, while DB9 and DB10 include only homogenous deformations, the



**Fig. 3 Main steps of the active learning procedure.** **a** Workflow of the active learning algorithm. (I) Fracture simulation by *K*-test at  $T = 0$  K. (II) Evaluation of the GAP predicted uncertainty for each frame generated during *K*-test. (III) Detection of extrapolation by comparing the maximum model uncertainty of each configuration with the predefined criterion. (IV) Construction of the crack-tip cell that is computable by DFT. (V) DFT calculation of the crack-tip cell. (VI) Refitting GAP. Steps (I)–(VI) are repeated until convergence is achieved. **b** Construction of the crack-tip DFT cell. (I) Identification of the atom with the largest uncertainty ( $ID_0$ ) in the deformed configuration. (II) Selection of a group of atoms ( $ID_g$ ) in the square centred on  $ID_0$  in the undeformed state. (III) Identification of the crack tip in group  $ID_g$  in the deformed state. (IV) Duplication and rotation of the selected region. (V) Generation of DFT cell containing two crack tips by merging the rotated replica and the original one.

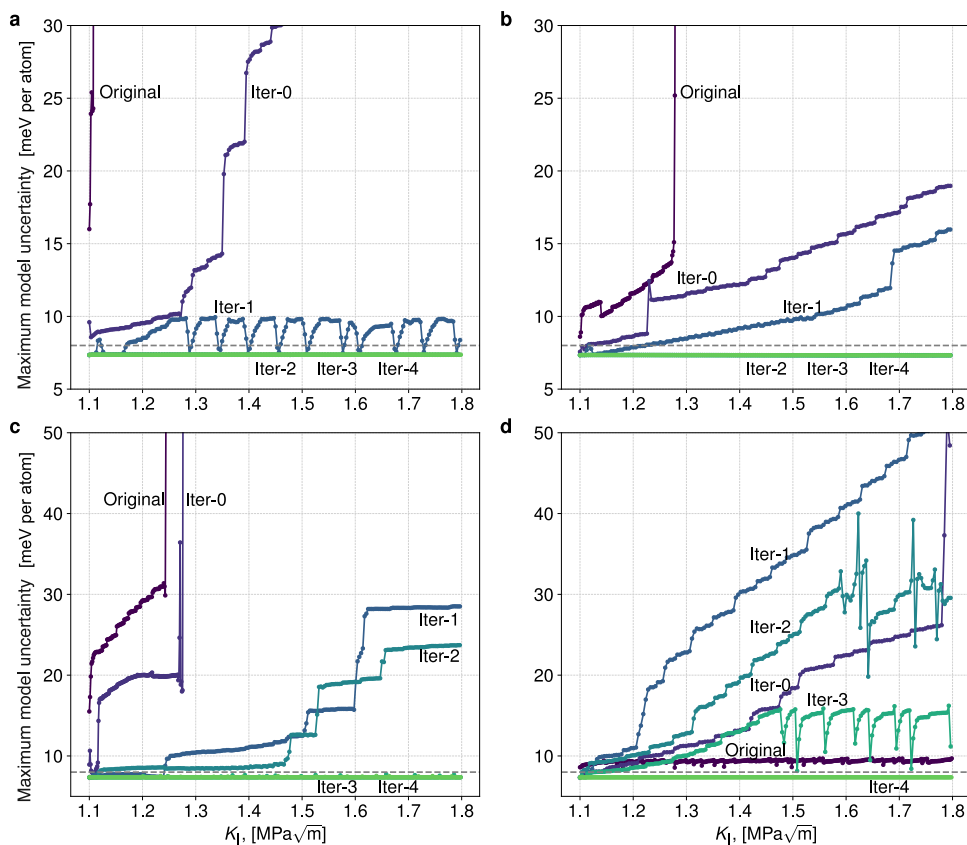
Step	Function	Code
I	<i>K</i> -test	LAMMPS <sup>51</sup>
II	Evaluation of the GAP predicted variance	QUIP <sup>48</sup>
III	Detection of extrapolation	Python
IV	Construction of the DFT cell	Python
V	DFT calculation	Quantum Espresso <sup>45</sup>
VI	Fitting GAP model	QUIP <sup>48</sup>

Roman numbers are consistent with Fig. 3a.

"*Iter-0*" GAP is not well-informed about crack tip local atomic environments. By virtue of the localisation assumption that the atom only interacts with its neighbouring atoms within a finite cutoff radius (5 Å in this study), the crack tip atomic environment can be represented by a computable DFT cell that contains the atomistic crack tip configuration. To learn efficiently from the crack tip, we develop an active learning algorithm by extracting crack tips from extrapolating configurations. The algorithm (see *Data Availability*) automatically identifies the extrapolated atomic

environments with respect to a predefined criterion and constructs a periodic crack cell that is computable with plane-wave DFT while minimising spurious surface effects. As a measure of the extrapolating degree from the training dataset, the square root of GAP predicted variance is used to assess the model uncertainty. The model uncertainty of new configurations generated during the simulation is evaluated and compared with a threshold. We consider the model to be saturated if the maximum model uncertainty does not exceed the threshold during the entire fracture simulation. The true error of the model, i.e., error with respect to DFT, is also evaluated to benchmark the accuracy of the converged model uncertainty. Since the lower-limit model uncertainty of Fe-GAP22 is ~6 meV per atom, we define the uncertainty threshold to be 10 meV per atom in this study.

Figure 3a shows the workflow, schematically indicating the main steps of the active learning algorithm. We first run a *K*-test and evaluate the GAP predicted variance for all frames dumped during the fracture simulation. Then, the maximum model uncertainty is compared with the predefined extrapolation criterion. A periodic cell is subsequently constructed by extracting the atomic configuration from the extrapolating frame. DFT calculations of the periodic crack-tip cell are performed, and the new DFT data are added to the training database. Next, GAP is retrained with the new DFT database. We repeat this process until



**Fig. 4** Maximum model uncertainty as a function of  $K$  for four crack systems. **(a)** (100)[011], **(b)** (110)[001], **(c)** (100)[010] and **(d)** (110)[ $\bar{1}\bar{1}0$ ]. The dashed line indicates the 8 meV per atom. *Original* is the GAP trained on the original database<sup>32</sup> (see Fig. 2 for the value of error). *Iter-0* is the GAP trained with original and preliminary database.

convergence is achieved, i.e., the maximum GAP predicted uncertainty is less than the predefined criterion during the entire fracture process. A summary of the software/code used in the active learning scheme is listed in Table 1.

One of the main challenges is the construction of an appropriate DFT cell that contains the crack tips at step (IV) of Fig. 3a. It is necessary to prevent spurious boundaries to avoid learning irrelevant/artificial atomic environments. The active learning has been applied to study the screw dislocation in bcc tungsten, in which the effective extrapolative configuration is constructed by taking advantage of symmetry<sup>35</sup>. The idea here is to symmetrise the crack tip along the crack plane normal to construct an approximate periodic cell. The construction of the DFT crack configuration is illustrated in Fig. 3b. First we identify the atom ( $ID_0$ ) with the largest model uncertainty at the crack tip in the deformed configuration. Then, a group of atoms ( $ID_g$ ) in a square centred on  $ID_0$  is selected in the undeformed configuration; once identified, these atoms form a crack tip configuration in the deformed state. Another crack tip configuration is created by duplicating the original one and rotating it by 180°. We construct the periodic DFT configuration by aligning the rotated replica and the original one by a rigid displacement. The periodic DFT configuration is surface-free and preserves the local atomic environment of the atom with the largest model uncertainty. Such construction allows both efficient DFT calculations and learning speed, as illustrated in the following section.

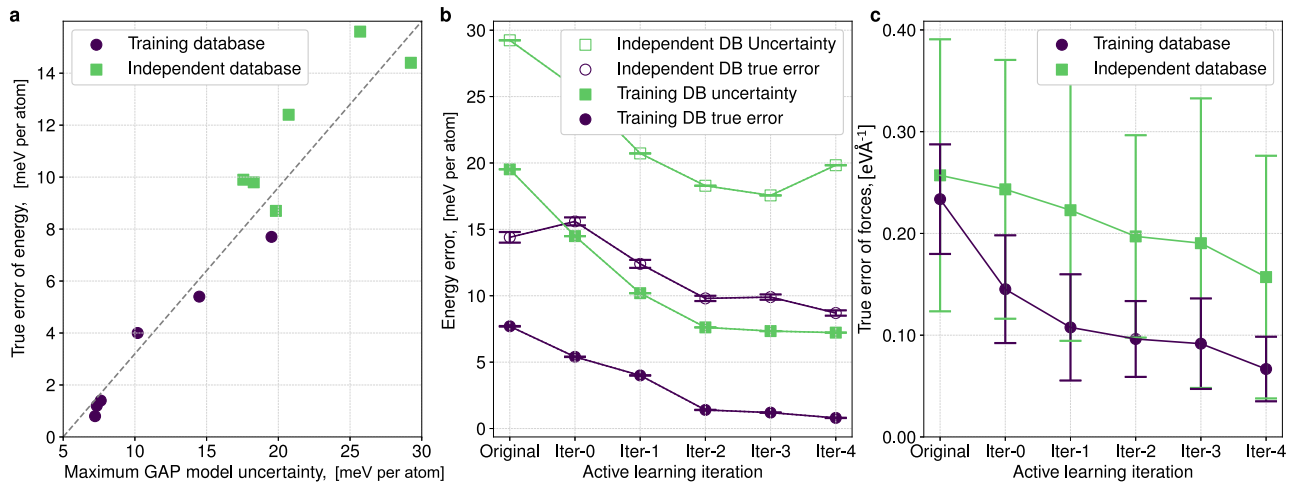
### Implementation of active learning

We implemented the active learning algorithm with the original Smooth Overlap of Atomic Positions (SOAP)<sup>36</sup> and an optimised SOAP descriptor<sup>37</sup> (referred to as “Turbo SOAP”). The presentation

here focuses on active learning results with Turbo SOAP because this leads to higher computing speed and faster convergence. The GAP model based on the original SOAP are discussed in Supplementary Fig. 4 and Supplementary Fig. 5, referred to as Fe-S-GAP22. We applied the algorithm to two crack systems in parallel, i.e., (100)[011] and (110)[001], such that GAP can learn the local environments from both {100} and {110} crack planes at the same time. Two sets of small crack tip regions are used, i.e., configurations extracted from fracture simulations with and without Gaussian noise ( $\mathcal{N} \sim [0, 0.05a_0]$ ) added to the three Cartesian coordinates. Figure 4 shows the convergence of the maximum model uncertainty at multiple  $K_I$ 's with respect to active learning iterations. The GAP model uncertainty is plotted as a function of  $K_I$  during a mode-I  $K$ -test at  $T = 0$  K.

As shown in Fig. 4a and b, the maximum model uncertainty converges to 8 meV per atom for crack systems (100)[011] and (110)[001] after two iterations of active learning. The maximum model uncertainty of (100)[010] converges to 8 meV per atom at the third iteration (Fig. 4c) while (100)[011] and (110)[001] crack systems remain of the same accuracy. As shown in Fig. 4d, the maximum model uncertainty of (110)[ $\bar{1}\bar{1}0$ ] increases after two iterations (*Iter-0* and *Iter-1*) and then converges after another three iterations. This is because the original GAP is able to extrapolate the atomic environments that are close to crack tip of (110)[ $\bar{1}\bar{1}0$ ]. However, the extrapolation is unstable and is easily deteriorated by adding more data, which means that the prediction is not fully converged yet. We added two crack-tips of (110)[ $\bar{1}\bar{1}0$ ] for one additional iteration of training and converged the maximum model uncertainty to 8 meV per atom (Fig. 4d).

To demonstrate the robustness of GAP uncertainty as an index of extrapolation, we evaluated the true error of two sets of DFT



**Fig. 5 Energy/force error and uncertainty analysis of the crack-tip configurations.** **a** The correlation between maximum model uncertainty and true energy error for crack-tip configurations from the DFT database of Fe-GAP22 (14 configurations) and crack-tip configurations from an independent DFT database (Fe-S-GAP22, 32 configurations). **b** Maximum model uncertainty and true error as functions of active learning iterations for two sets of DFT crack-tip configurations. **c** Force error of crack-tip atoms as functions of active learning iterations. The error bars indicate the standard deviation, that is computed based on all crack-tip configurations.

configurations, i.e., the crack-tip configurations from the database of Fe-GAP22 and a set of independent crack-tip configurations from Fe-S-GAP22. The true errors of energy and forces are computed for a batch of potentials in Fig. 5. Figure 5a plots the correlation between maximum GAP uncertainty and true error per-atom, indicating a quasi-linear relation. Both maximum GAP uncertainty and true error converge with the active iterations, displaying the same trend (Fig. 5b). Since the largest model uncertainty always locates at crack-tip atoms, we evaluated the true error of forces on the crack-tip atoms (Fig. 5b). The true force error decreases to 0.1 and 0.2 eV Å<sup>-1</sup> for training and independent crack-tip configurations, showing the accuracy of the model.

### Prediction of fracture mechanisms and critical stress intensity factor ( $K_{Ic}$ )

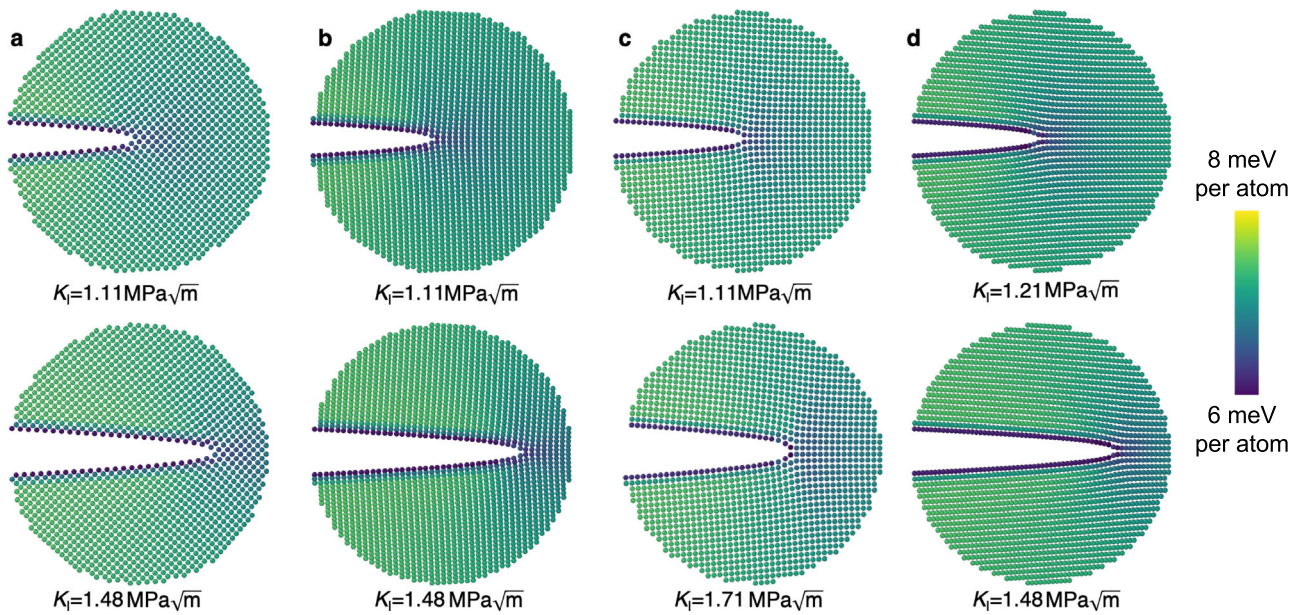
We predicted the lattice constant, elastic constants, surface energies and generalized stacking fault energy (GSFE) profiles as a benchmark of the Fe-GAP22, Fe-S-GAP22 and found similar accuracy as Fe-GAP18 (see Supplementary Figs. 6 and 7). Subsequently, we performed fracture simulation of single-crystal ferromagnetic iron at  $T=0$  K and finite temperatures ( $T=1, 10, 100, 200, 300$  K) under high loading rate ( $10^9$  MPa $\sqrt{m}$ s<sup>-1</sup>). To investigate the uncertainty effects of the Gaussian process, we trained 10 different GAPs and performed  $K$ -test on (100)[011] crack system based on those potentials (see Supplementary Fig. 8). We found that the  $K_{Ic}$  predicted by different models varies within  $\pm 0.005$  MPa $\sqrt{m}$ , which is the accuracy limit introduced by the uncertainty of the ML algorithm. Therefore, 0.01 MPa $\sqrt{m}$  is used as the incremental step during the  $K$ -test. The fracture mechanism at  $T=0$  K is found to be cleavage on the original plane in all considered cases with a converged GAP model uncertainty during the entire cleavage process (Fig. 6). Neither phase transformation nor {110} planar faults appear according to Fe-GAP22 and Fe-S-GAP22 predictions.

The cleavage is always taking place on the plane that has the maximum normal stress. To evaluate the normal stress, we performed rigid body separation of {100} and {110} planes by using DFT and Fe-GAP22 to calculate the universal binding energy relation (UBER) (Fig. 7a). Fe-GAP22 predicted UBER of {100} plane perfectly overlaps with DFT ones, while the Fe-GAP22 predicted UBER of {110} plane slightly deviates from DFT calculations. We further computed the normal stress during the separation (Fig. 7b) by taking the derivative of the fitted curve in Fig. 7a. The minimum

stress required to separate {100} and {110} surfaces predicted by DFT are 32.68 and 34.62 GPa, respectively, and Fe-GAP22 predictions are 0.3% and 9.7% larger than DFT results. For all crack systems considered here, the deviation of the crack from the original plane to {100} or {110} is not likely to occur because it requires a lower cohesive strength than on the original plane. The Fe-GAP22 prediction of cohesive strength confirms that the cleavage will take place on the original crack plane, which is also consistent with anisotropic CLE calculations, i.e., the largest normal stress exists on the original plane under pure mode-I loading (see Supplementary Fig. 9).

At finite temperatures, cleavage is consistently taking place on the original plane for all crack systems, which shows that cleavage is the controlling mechanism of fracture in iron at high loading rates when considering finite temperatures ( $T \leq 300$  K). At  $T=300$  K, the maximum model uncertainty raises to  $\sim 30$  meV per atom for (110)[1 $\bar{1}$ 0] and  $\sim 10$  meV per atom for the other three crack systems (see Supplementary Fig. 10). Five independent  $K$ -tests are performed at each temperature for the statistic prediction of  $K_{Ic}$  at finite temperatures, as indicated by the error bar shown in Fig. 8. Here we propose a criterion to determine  $K_{Ic}$  for cleavage from atomistic simulation results based on the traction-separation curve. The first pair of atoms at the crack tip can still interact with each other after the initial debonding process until the entire separation is achieved, i.e., the distance between them reaches the cutoff distance in atomistic simulations. Therefore,  $K_{Ic}$  is defined as the point where the crack tip bond is separated to a cutoff distance and will remain open upon further increasing  $K_I$ . Inspection of the atomic snapshots reveals that the fracture mechanism for all crack systems remains cleavage up to  $T=300$  K as predicted by Fe-S-GAP22, which demonstrates that the predicted cleavage mechanism is independent of the specific crack-tip converged DFT database.

Fe-GAP22 predicted  $K_{Ic}$  for all crack systems decreases from 0 K to 1 K and converges from  $T=1$  K to 200 K for all crack systems except for (110)[001] (Fig. 8). For this crack system, larger lattice trapping is predicted at  $T=0$  K by evaluating lattice trapping range  $\Delta K$  according to refs. 38,39. The ratio between the cohesive strength and the maximum interaction range ( $\sigma_{coh}/\delta_{sep}^{max}$ ) is used to quantify the lattice trapping in tungsten, which demonstrates a quasi-linear relation<sup>40</sup>. However, no clear linear dependency is found in the case of bcc iron (see Supplementary Fig. 11). As shown by Mak et al.<sup>4</sup>, lattice trapping is associated with a crack



**Fig. 6** Atomic snapshots showing the fracture mechanism at  $T=0\text{K}$  predicted by Fe-GAP22. **a** (100)[010], **(b)** (100)[011], **(c)** (110)[001], and **(d)** (110)[110]. The configuration is taken at  $K_I$  given below each snapshot. Atoms are coloured according to the GAP model uncertainty.

propagation activation energy barrier that vanishes if  $K_{Ic} = K_G$ . Since the difference between  $K_{Ic}$  and  $K_G$  for (110)[001] drops significantly with short MD at low temperature ( $T = 1\text{--}100\text{K}$ ), a small lattice trapping activation energy barrier is expected to account for the discrepancy between  $K_{Ic}$  at  $T = 0\text{K}$  and  $100\text{K}$ . As plotted in Fig. 8b, Fe-S-GAP22 predicted  $K_{Ic}$  has almost no difference between  $T = 0\text{K}$  and  $100\text{K}$  (see Supplementary Fig. 12 for prediction of the other 3 crack systems). Therefore, we conclude that the difference between  $K_{Ic}$  at  $T = 0\text{K}$  and  $100\text{K}$  predicted by Fe-GAP22 is a consequence of the rougher PES predicted by Fe-GAP22 compared with Fe-S-GAP22. However, the small activation barrier does not change the fundamental cleavage mechanism. The predicted  $K_{Ic}$ 's of Fe-GAP22 and Fe-S-GAP22 converge to the same value within a  $0.004\text{MPa}\sqrt{m}$  error bound. All predictions at  $T > 10\text{K}$  are within 10% of Griffith prediction ( $K_G$ ), indicating mild lattice trapping effects<sup>41</sup>. At high temperatures ( $T > 100\text{K}$ ), the surfaces near the crack tip can be closed due to thermal fluctuations, which leads to a slight increase of  $K_{Ic}$  at  $T = 200\text{K}$  and  $300\text{K}$  for all crack systems.

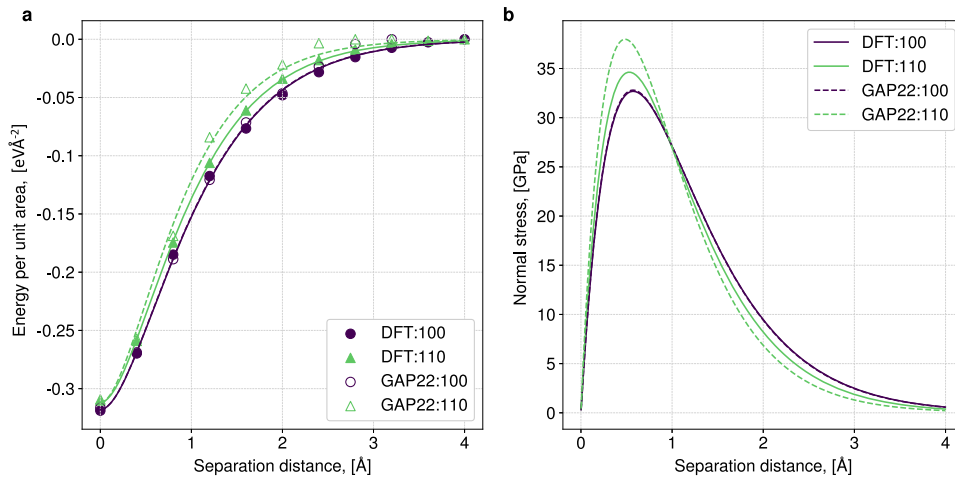
## DISCUSSION

Our analysis unveils that the  $T = 0\text{K}$  fracture mechanism for {100} and {110} crack planes is cleavage on the pre-existing crack planes. There are no fracture experiments performed close to  $T = 0\text{K}$ , thus qualitative comparison is possible only with the work of ref. 5, which is performed at  $T = 77\text{K}$  and finite loading rates that are inaccessible to direct MD simulations. The prediction of Fe-GAP22 qualitatively agrees with experiments on {100} crack plane, i.e., cleavage on the original crack plane. This is a considerable step forward with respect to all classical potentials considered by Möller and Bitzek<sup>15</sup>, and in the introduction of the present paper, that, in contrast with Fe-GAP22, cannot predict pure cleavage on {100} plane for (100)[010] crack system. For example, most of the classical potentials predict phase transition at the crack-tip in (100)[010] crack system since the empirical potentials may not be able to reproduce the ground state energy of different phases (bcc, fcc, hcp) of iron<sup>15</sup>. Fe-GAP22 is constructed using a consistent ferromagnetic DFT database, which prevents the artificial phase transition at the highly strained region. Moreover, Fe-GAP22 is free from {110} planar faults that

are caused by the artificial minimum of the strain-dependent GSFE predicted by classical potentials<sup>42</sup>. However, in the case of {110} crack planes, Fe-GAP22 also predicts cleavage on the original plane while experiments show crack kinking into {100} planes<sup>5</sup>. Our calculation of UBER by Fe-GAP22 and anisotropic CLE analysis shows that, under  $K_I$  loading and in the absence of other defects close to the crack tip, cleavage should take place on the original crack plane for both {100} and {110}. The experiments, instead, show a semi-brittle fracture behaviour of bcc iron because cleavage is accompanied by extensive dislocation activity<sup>5</sup>.

Therefore, in order to perform quantitative connection to experiments, the temperature and loading-rate dependent competition between cleavage and dislocation emission must be assessed, e.g., by performing Nudged Elastic Band (NEB) calculations and using transition state theory<sup>4,43</sup>. Also, the interaction between pre-existing defects (e.g., nanovoids, dislocations, etc) and cracks is expected to change the stress field around the crack tip and hence affect the crack-tip response. In the present atomistic fracture simulations however, the specimen is defect-free (except for the sharp crack), and the boundary conditions are those of a half-infinite crystal in a 2D setup. Therefore, quantitative predictions of fracture require a multiscale approach. Despite the fact that a quantitative comparison of  $K_{Ic}$  is not possible, we extrapolate the experimental  $K_{Ic}$  linearly to  $T = 0\text{K}$ , i.e.,  $K_{Ic} = 3.7, 3.8, 7.3,$  and  $11.6\text{MPa}\sqrt{m}$ , for (100)[010], (100)[011], (110)[001], and (110)[110] crack systems, respectively<sup>5</sup>.  $K_{Ic}$  for {110} crack plane is 2–3 times larger than {100}, suggesting pronounced thermally-activated dislocation plasticity at finite temperatures on {110} crack plane. Such a significant energy dissipation is expected to account for the crack deviation from {110} to {100} crack plane. Thus, large-scale simulations in which rate and temperature dependence is considered along with pre-existing defects should be performed to enable prediction of the experimental  $K_{Ic}$ .

Griffith<sup>24</sup> ( $K_G$ ) and Rice<sup>2</sup> ( $K_{Ic}$ ) theories are used to predict cleavage and dislocation emission from the crack tip. Since cleavage is the controlling fracture mechanism in bcc iron at  $T = 0\text{K}$ , we compare the Fe-GAP22 predicted  $K_{Ic}$  with Griffith theory  $K_G$ . Because  $K_G$  only accounts for the energy of newly created surface, it serves as a lower bound of  $K_{Ic}$ . Thus, the excess part of Fe-GAP22 predicted  $K_{Ic}$  compared with  $K_G$  could be used as an



**Fig. 7** Energy change and traction as a function of the rigid separation process. **a** Universal binding energy relation (UBER) for {100} and {110} planes<sup>53</sup>. Filled and unfilled symbols represent DFT and Fe-GAP22 predictions respectively. Circle and triangle symbols indicate {100} and {110} planes respectively. Solid and dash lines are DFT and Fe-GAP22 results fitted to UBER equation. **b** Normal stress on {100} and {110} planes by taking derivative of fitted curves from (a).

indicator of lattice trapping effects. For all three crack systems except for (110)[ $\bar{1}\bar{1}0$ ], the predicted  $K_{Ic}$  is around 10% larger than Griffith criterion, which is also found in bcc high entropy alloys<sup>4</sup>. (110)[ $\bar{1}\bar{1}0$ ] is 5% larger than  $K_G$ , indicating less lattice trapping effects.

We obtained a Fe GAP that is able to predict atomistic fracture in iron with a converged model uncertainty based on an existing database, showing that GAP can be improved for fracture study. The current approach is based on two essential ingredients, i.e., fracture-relevant configurations and active-learned crack-tip configurations. Since the model uncertainty and true error decrease significantly and converge in a few active learning iterations (Figs. 4 and 5) by adding crack-tip configurations, we conclude that the improved fracture predictability of Fe-GAP22, compared to Fe-GAP18, is enabled by the extended DFT database, especially crack-tip configurations. Here, we have shown that a general-purpose DFT database cannot guarantee fracture predictability in the case of bcc iron and the active learning strategy brings the possibility of developing potentials for specific applications. In particular, active learning enables efficient learning of the extrapolated local atomic environments that are relevant to fracture, which requires only a few DFT calculations. Our strategy can be generalised to other materials and other ML frameworks, providing a systematic way to improve ML potentials for studying fracture behaviour, in line with active learning approaches.

For all crack systems investigated here, we found that crack cleaves on the pre-existing crack plane. Based on the convergence of Fe-GAP22 predicted  $K_{Ic}$  with respect to temperatures, we conclude that the increase of  $K_{Ic}$  at  $T=0$  K is caused by minute barriers. The predicted  $K_{Ic}$  at  $T=0$  K is converged with respect to local symmetry breaking, which we have studied by performing high loading rate mode-I tests from  $T=10$  K up to  $T=300$  K. Our results confirm the  $T=0$  K atomistic fracture mechanism in ferromagnetic iron under mode-I loading, settling the inconsistency of crack-tip mechanisms based on classical potentials.

We conclude by pointing out that Fe-GAP22 can describe cracks and screw dislocations with quantum accuracy<sup>26</sup>. It should be noted that our fracture simulation is based on a quasi-2D setup, which enables a direct comparison with previous studies<sup>14,15</sup>. Recently, a numerical continuation method based on flexible boundary conditions has been proposed, which allows accurate prediction of the lattice trapping barriers<sup>44</sup>. The energy barriers associated with the fracture-related processes (crack propagation and dislocation emission) are currently being investigated by NEB

and numerical continuation method along with Fe-GAP22. Based on transition state theory, the competition of atomistic crack-tip mechanisms (cleavage and dislocation emission) can be connected to microscale behaviour to assess the brittle to ductile transition at finite temperatures and finite loading rates<sup>4</sup>. Mobility laws for microscale simulations can thus be developed by using Fe-GAP22.

## METHODS

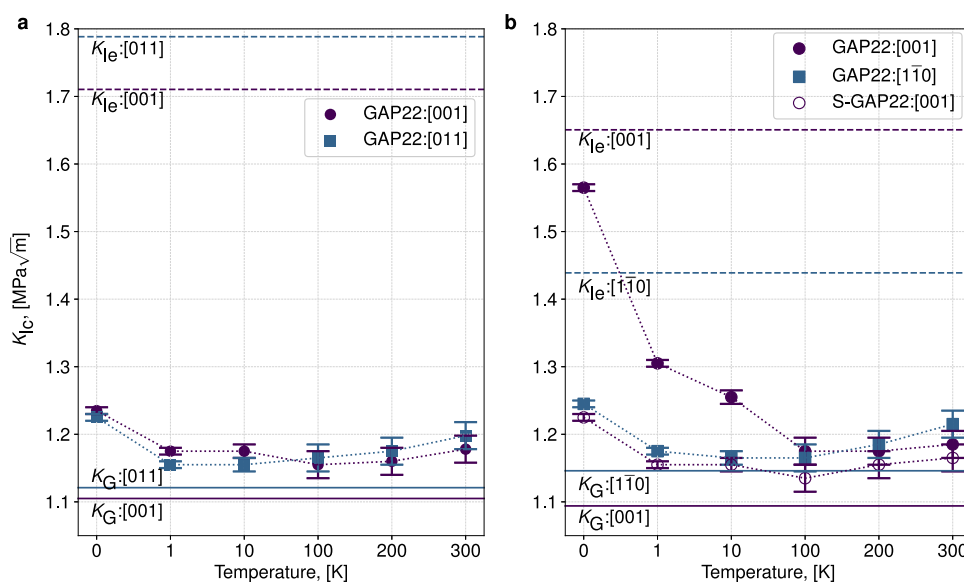
### DFT calculation

All DFT calculations were performed in collinear spin-polarised plane wave method implemented within QUANTUM ESPRESSO<sup>45</sup>. An ultrasoft GGA PBE pseudopotential from 0.021pslibrary is employed<sup>46</sup>. The kinetic energy cutoff for wavefunctions and charge density are set to 90 Ry and 1080 Ry, respectively. The Brillouin Zone is integrated with a Monkhorst-Pack grid and a Marzari-Vanderbilt smearing scheme at an effective temperature of 0.01 Ry<sup>47</sup>. The parameter to initialise the magnetisation is set to 0.34, which is a fractional parameter between  $-1$  (all spins down for the valence electrons) to  $1$  (all spins up). An adaptive  $k$ -mesh approach is used to calculate primitive cells (DB9 and DB10), ensuring that the  $k$ -spacing is smaller than  $0.015 \text{ \AA}^{-1}$ . For other calculations, the simulation cell size is chosen to satisfy the condition that  $k$  spacing is smaller than  $0.03 \text{ \AA}^{-1}$  while only the  $\Gamma$  point is used along the separation direction. A benchmark study compared with Dragoni's database is conducted to ensure all the parameter settings can reach the convergence of 1 meV per atom,  $0.01 \text{ eV \AA}^{-1}$  and 0.01 GPa for energies, forces and stresses, respectively.

### GAP training

A parallel version (1645177290) of the open-source package QUANTUM mechanics and Interatomic Potentials (QUIP) is used to fit the GAP model<sup>48</sup>. We used three descriptors to encode the local atomic environment, i.e., one distance-based two-body interaction and two many-body turbo-SOAP descriptors. Two turbo-SOAP descriptors are used to describe the inner and outer atomic environment within the cutoff radius of  $3.5 \text{ \AA}$  and  $5 \text{ \AA}$ , respectively. Dot product kernel is used for both turbo-SOAP descriptors, and the CUR matrix decomposition procedure is applied to find the optimal representative local atomic environments. A power of 4 is added to the dot product to sharpen the sensitivity of changing the atomic positions. The cutoff smoothing distance is





**Fig. 8** Critical  $K_I$  predicted by Fe-GAP22 at different temperatures ( $T=0\text{K} - 300\text{K}$ ). **a** (100) and **(b)** (110) crack plane. The solid and dashed line indicate critical  $K$  predicted by Griffith ( $K_G$ ) and Rice ( $K_{Ie}$ ) theories, respectively. Fe-S-GAP22 is the GAP trained on the active learning database developed based on original SOAP. The error bars indicate the standard deviation, that is computed based on five independent realisations.

set to  $1 \text{ \AA}$ . The number of radial and angular basis for turbo-SOAP are set to 8. The number of representative points is 4700. The default regularisation for energies, forces and stresses are  $0.005 \text{ eV}$ ,  $0.2 \text{ eV \AA}^{-1}$  and  $0.01 \text{ eV \AA}^{-3}$ . The training command line is available in *Data Availability*.

### MS/MD simulation setup

We used a cylinder-shaped half crack setup (see schematic plot in Supplementary Fig. 13), where  $x$ ,  $y$  and  $z$  are aligned with the crack propagation direction, crack plane normal and crack front, respectively. The diameter of the model in  $x$ - $y$  plane is set to  $300 \text{ \AA}$ , which satisfies the requirement that the fracture process zone is much smaller than the simulation cell, as proved in ref. <sup>3,15</sup>. A convergence analysis is conducted to ensure the predicted  $K_{Ic}$  is size-independent, which is shown in Supplementary Fig. 14. The boundary of the simulation cell ( $K$ -dominant region) is governed by linear elasticity, ensuring a meaningful prediction of ( $K_{Ic}$ ). It has been shown that the opening stress and predicted  $K_{Ic}$  are converged with a cell size of  $300 \text{ \AA}^3$ . We considered plane strain conditions in a 2D setting by imposing periodic boundary conditions along the crack front direction.

We used a  $K$ -test loading scheme that allows controlling the fracture process by directly increasing the stress intensity factor  $K_I$ .  $K$ -test implements a displacement-controlled loading scheme, which makes use of anisotropic CLE. The displacement field of a half crack in an infinite anisotropic medium can be derived via Lekhnitskii's formalism<sup>49</sup>, as summarised in Supplementary Eqs. 1–8. The asymptotic stress field near the crack tip can be uniquely characterised by the stress intensity factor  $K_I$ , which enables a single-parameter controlled loading scheme<sup>3</sup>. The  $K_I$  displacement field can then be applied to the boundary of the simulation cell directly. During  $K$ -test,  $K_I$  is increased with a step of  $\Delta K_I = 0.01 \text{ MPa}\sqrt{\text{m}}$ . The displacement is applied incrementally at each step, and the system is equilibrated under the constraining of fixed boundary atoms. In the MS analysis, conjugate gradient (CG) with a force tolerance of  $10^{-9} \text{ eV \AA}^{-1}$  is first applied to minimise the potential energy of the system. Next, FIRE<sup>50</sup> is applied with a force tolerance of  $10^{-3} \text{ eV \AA}^{-1}$  to drive the system away from local minima, which allows the system to achieve better convergence. For MD, we used the Nosé-Hoover thermostat

with a timestep of  $0.001 \text{ ps}$  and the system is equilibrated for  $10 \text{ ps}$  at each incremental step.

To preserve an atomistically sharp crack, we neither screen the interaction between free crack surfaces nor delete any atoms (so-called screening and blunting<sup>3</sup>). Instead, we started with a  $K_{init}$  that maintains the current crack tip position. However, it should be noted that finding  $K_{init}$  may require trial-and-error. All molecular statics/dynamics calculations in this work were performed by using LAMMPS<sup>51</sup>. The atomic configurations are visualised in OVITO<sup>52</sup>.

### DATA AVAILABILITY

The training command line and extended DFT database are available on Materials Cloud <https://archive.materialscloud.org/record/2022.102>.

### CODE AVAILABILITY

The active learning algorithm is available on our Github page <https://github.com/leiapple/Fe-GAP22-AL>.

Received: 8 August 2022; Accepted: 24 November 2023;  
Published online: 08 December 2023

### REFERENCES

- Hartmaier, A. & Gumbsch, P. Scaling relations for crack-tip plasticity. *Philos. Mag. A* **82**, 3187–3200 (2002).
- Rice, J. R. Dislocation nucleation from a crack tip: an analysis based on the Peierls concept. *J. Mech. Phys. Solids* **40**, 239–271 (1992).
- Andric, P. & Curtin, W. A. Atomistic modeling of fracture. *Model. Simul. Mater. Sci. Eng.* **27**, 013001 (2018).
- Mak, E., Yin, B. & Curtin, W. A. A ductility criterion for bcc high entropy alloys. *J. Mech. Phys. Solids* **152**, 104389 (2021).
- Hribernik, M. L., *Cleavage Oriented Iron Single Crystal Fracture Toughness* PhD thesis (University of California, Santa Barbara, 2006).
- deCelis, B., Argon, A. S. & Yip, S. Molecular dynamics simulation of crack tip processes in alpha-iron and copper. *J. Appl. Phys.* **54**, 4864–4878 (1983).
- Guo, Y.-F., Wang, Y.-S. & Zhao, D.-L. Atomistic simulation of stress-induced phase transformation and recrystallization at the crack tip in bcc iron. *Acta Mater.* **55**, 401–407 (2007).
- Guo, Y.-F. & Gao, Y.-C. Combined atomistic simulation and continuum mechanics: size-dependent behavior of atomistic simulation for brittle fracture in bcc-iron. *Comp. Mater. Sci.* **36**, 432–439 (2006).

9. Guo, Y.-F. & Zhao, D.-L. Atomistic simulation of structure evolution at a crack tip in bcc-iron. *Mater. Sci. Eng. A* **448**, 281–286 (2007).
10. Cao, L.-X. & Wang, C.-Y. Atomistic simulation for configuration evolution and energetic calculation of crack in body-centered-cubic iron. *J. Mater. Res.* **21**, 2542–2549 (2006).
11. Wang, Z. et al. Atomistic simulation of martensitic transformations induced by deformation of  $\alpha$ -Fe single crystal during the mode-I fracture. *J. Mater. Sci.* **56**, 1–21 (2021).
12. Meiser, J. & Urbassek, H. M.  $\alpha \leftrightarrow \gamma$  phase transformation in iron: comparative study of the influence of the interatomic interaction potential. *Model. Simul. Mater. Sci. Eng.* **28**, 055011 (2020).
13. Cheung, K. S. & Yip, S. A molecular-dynamics simulation of crack-tip extension: the brittle-to-ductile transition. *Model. Simul. Mater. Sci. Eng.* **2**, 865–892 (1994).
14. Gordon, P., Neeraj, T., Luton, M. J. & Farkas, D. Crack-Tip deformation mechanisms in  $\alpha$ -Fe and Binary Fe alloys: an atomistic study on single crystals. *Metall. Mater. Trans. A* **38**, 2191–2202 (2007).
15. Möller, J. J. & Bitzek, E. Comparative study of embedded atom potentials for atomistic simulations of fracture in  $\alpha$ -iron. *Model. Simul. Mater. Sci. Eng.* **22**, 045002 (2014).
16. Kermode, J. et al. Low-speed fracture instabilities in a brittle crystal. *Nature* **455**, 1224–1227 (2008).
17. Zhou, X. W., Johnson, R. A. & Wadley, H. N. G. Misfit-energy-increasing dislocations in vapor-deposited CoFe/NiFe multilayers. *Phys. Rev. B* **69**, 144113 (2004).
18. Olsson, P. A. Semi-empirical atomistic study of point defect properties in BCC transition metals. *Comp. Mater. Sci.* **47**, 135–145 (2009).
19. Liyanage, L. S. I. et al. Structural, elastic, and thermal properties of cementite ( $\text{Fe}_3\text{C}$ ) calculated using a modified embedded atom method. *Phys. Rev. B* **89**, 094102 (2014).
20. Asadi, E., Asle Zaeem, M., Nouranian, S. & Baskes, M. I. Quantitative modeling of the equilibration of two-phase solid-liquid Fe by atomistic simulations on diffusive time scales. *Phys. Rev. B* **91**, 024105 (2015).
21. Etesami, S. A. & Asadi, E. Molecular dynamics for near melting temperatures simulations of metals using modified embedded-atom method. *J. Phys. Chem. Solids* **112**, 61–72 (2018).
22. Byggmästar, J. & Granberg, F. Dynamical stability of radiation-induced C15 clusters in iron. *J. Nucl. Mater.* **528**, 151893 (2020).
23. Starikov, S. et al. Angular-dependent interatomic potential for large-scale atomistic simulation of iron: development and comprehensive comparison with existing interatomic models. *Phys. Rev. Mater.* **5**, 063607 (2021).
24. Griffith, A. A. The phenomena of rupture and flow in solids. *Philos. Trans. R. Soc. A* **221**, 163–198 (1921).
25. Sun, Y. & Beltz, G. E. Dislocation nucleation from a crack tip: a formulation based on anisotropic elasticity. *J. Mech. Phys. Solids* **42**, 1905–1932 (1994).
26. Maresca, F., Dragoni, D., Csányi, G., Marzari, N. & Curtin, W. A. Screw dislocation structure and mobility in body centered cubic Fe predicted by a Gaussian Approximation Potential. *npj Comp. Mater.* **4**, 1–7 (2018).
27. Bartók, A. P., Kermode, J., Bernstein, N. & Csányi, G. Machine learning a general-purpose interatomic potential for silicon. *Phys. Rev. X* **8**, 041048 (2018).
28. Bartók, A. P. & Csányi, G. Gaussian approximation potentials: a brief tutorial introduction. *Int. J. Quantum Chem.* **115**, 1051–1057 (2015).
29. Jinnouchi, R., Lahnsteiner, J., Karsai, F., Kresse, G. & Bokdam, M. Phase transitions of hybrid perovskites simulated by machine-learning force fields trained on the fly with bayesian inference. *Phys. Rev. Lett.* **122**, 225701 (2019).
30. Jinnouchi, R., Karsai, F. & Kresse, G. On-the-fly machine learning force field generation: application to melting points. *Phys. Rev. B* **100**, 014105 (2019).
31. Dragoni, D., Daff, T. D., Csányi, G. & Marzari, N. Achieving DFT accuracy with a machine-learning interatomic potential: thermomechanics and defects in bcc ferromagnetic iron. *Phys. Rev. Mater.* **2**, 013808 (2018).
32. Dragoni, D., Daff, T. D., Csányi, G. & Marzari, N., Gaussian approximation potentials for iron from extended first-principles database (Data Download). *Materials Cloud Archive*. **2017.0006/v2**, <https://doi.org/10.24435/materialscloud:2017.0006/v2> (2017).
33. Sobol, I. M. On the distribution of points in a cube and the approximate evaluation of integrals. *J. Comput. Math. Math. Phys.* **7**, 784–802 (1967).
34. Andric, P. & Curtin, W. A. New theory for Mode I crack-tip dislocation emission. *J. Mech. Phys. Solids* **106**, 315–337 (2017).
35. Hodapp, M. & Shapeev, A. In operando active learning of interatomic interaction during large-scale simulations. *Mach. Learn.: Sci. Technol.* **1**, 045005 (2020).
36. Bartók, A. P., Kondor, R. & Csányi, G. On representing chemical environments. *Phys. Rev. B* **87**, 184115 (2013).
37. Caro, M. A. Optimizing many-body atomic descriptors for enhanced computational performance of machine learning based interatomic potentials. *Phys. Rev. B* **100**, 024112 (2019).
38. Fuller Jr, E. & Thomson, R. Nonlinear lattice theory of fracture in *Analysis and Mechanics*, 387–394 (Elsevier, 1978).
39. Möller, J. J. & Bitzek, E. Fracture toughness and bond trapping of grain boundary cracks. *Acta Mater.* **73**, 1–11 (2014).
40. Hiremath, J. P., Melin, S., Bitzek, E. & Olsson, P. A. Effects of interatomic potential on fracture behaviour in single-and bicrystalline tungsten. *Comp. Mater. Sci.* **207**, 111283 (2022).
41. Curtin, W. A. On lattice trapping of cracks. *J. Mater. Res.* **5**, 1549–1560 (1990).
42. Möller, J. J. et al. {110} planar faults in strained bcc metals: Origins and implications of a commonly observed artifact of classical potentials. *Phys. Rev. Mater.* **2**, 093606 (2018).
43. Zhu, T., Li, J. & Yip, S. Atomistic configurations and energetics of crack extension in silicon. *Phys. Rev. Lett.* **93**, 205504 (2004).
44. Buze, M. & Kermode, J. R. Numerical-continuation-enhanced flexible boundary condition scheme applied to mode-I and mode-III fracture. *Phys. Rev. E* **103**, 033002 (2021).
45. Giannozzi, P. et al. QUANTUM-ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Condens. Matter Phys.* **21**, 395502 (2009).
46. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
47. Marzari, N., Vanderbilt, D., De Vita, A. & Payne, M. C. Thermal contraction and disordering of the Al(110) Surface. *Phys. Rev. Lett.* **82**, 3296–3299 (1999).
48. Csányi, G. et al. Expressive programming for computational physics in Fortran 95+. *Newsletter Comput. Phys. Group*, 1–24 (2007).
49. Ting, T. C. T., *Anisotropic elasticity theory and applications 1st edn*, (Oxford University, 1996).
50. Bitzek, E., Koskinen, P., Gähler, F., Moseler, M. & Gumbusch, P. Structural relaxation made simple. *Phys. Rev. Lett.* **97**, 170201 (2006).
51. Thompson, A. P. et al. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comp. Phys. Comm.* **271**, 108171 (2022).
52. Stukowski, A. Visualization and analysis of atomistic simulation data with OVITO—the open visualization tool. *Model. Simul. Mater. Sci. Eng.* **18**, 015012 (2009).
53. Rose, J. H., Smith, J. R. & Ferrante, J. Universal features of bonding in metals. *Phys. Rev. B* **28**, 1835–1845 (1983).

## ACKNOWLEDGEMENTS

This work made use of the Dutch national e-infrastructure with the support of the SURF Cooperative using grant no. EINF-2393 and EINF-3104. We thank the Centre for Information Technology of the University of Groningen (UG) for their support and for providing access to the Peregrine high performance computing cluster. L.Z. would like to thank Predrag Andric for LAMMPS implementation of fracture simulation and Miguel Caro for the implementation of Turbo SOAP descriptor. F.M. acknowledges the support from the start-up grant from the Faculty of Science and Engineering at the University of Groningen.

## AUTHOR CONTRIBUTIONS

All authors designed the research together. L.Z. and F.M. developed the active learning framework. L.Z. implemented the framework and performed all the DFT, MS and MD calculations. All authors analysed the data, discussed the results, wrote the manuscript together and contributed to the discussions and revisions of the paper.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41524-023-01174-6>.

**Correspondence** and requests for materials should be addressed to Lei Zhang or Francesco Maresca.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023