Accuracy and transferability of Gaussian approximation potential models for tungsten

Wojciech J. Szlachta, Albert P. Bartók, and Gábor Csányi
Engineering Laboratory, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, United Kingdom
(Received 16 May 2014; revised manuscript received 26 August 2014; published 24 September 2014)

We introduce interatomic potentials for tungsten in the bcc crystal phase and its defects within the Gaussian approximation potential framework, fitted to a database of first-principles density functional theory calculations. We investigate the performance of a sequence of models based on databases of increasing coverage in configuration space and showcase our strategy of choosing representative small unit cells to train models that predict properties observable only using thousands of atoms. The most comprehensive model is then used to calculate properties of the screw dislocation, including its structure, the Peierls barrier and the energetics of the vacancy-dislocation interaction. All software and raw data are available at www.libatoms.org.

DOI: 10.1103/PhysRevB.90.104108 PACS number(s): 65.40.De, 31.50.−x, 34.20.Cf, 71.15.Nc

Tungsten is a hard, refractory metal with the highest melting point (3695 K) among metals, and its alloys are utilized in numerous technological applications. The details of the atomistic processes behind the plastic behavior of tungsten have been investigated for a long time, and many interatomic potentials exist in the literature reflecting an evolution, over the past three decades, in their level of sophistication, starting with the Finnis-Sinclair (FS) potential [1], embedded atom model (EAM) [2], various other FS and EAM parametrizations [3–6], modified embedded atom models (MEAMs) [7–10], and bond order potentials (BOPs) [11–13]. While some of these methods have been used to study other transition metals [14–16], there is renewed interest in modeling tungsten due to its many high-temperature applications—e.g., it is one of the candidate materials for plasma facing components in the JET and ITER fusion projects [17–19].

A recurring problem with empirical potentials, due to the use of fixed functional forms with only a few adjustable parameters, is the lack of flexibility: when fitted to reproduce a given property, predictions for other properties can have large errors. Figure 1 shows the basic performance of the BOP and MEAM, two of the more sophisticated potentials that reproduce the correct screw dislocation core structure, and also the simpler FS potential,[1] all in comparison with results of density functional theory (DFT). While the figure emphasizes fractional accuracy, we show the corresponding absolute numerical values in Table I. The BOP is poor in describing the vacancy but is better at surfaces, whereas the MEAM is the other way around. While this compromise can sometimes be made with good judgment for specific applications, many interesting properties, particularly those that determine the material behavior at larger length scales, arise from the competition between different atomic-scale processes, which therefore all need to be described equally well. For example, dislocation pinning, depinning, and climb involve both elastic properties and core structure, as well as the interaction of dislocations with defects. Ways to deal with this problem include use of multiple levels of accuracy as in quantum mechanics/molecular mechanics [20] or allowing the parameters of the potential to vary in time and space [21].

Here we describe a milestone in a research program aimed at creating a potential that circumvents the problem of fixed functional forms. The purpose of the present work is twofold. First, we showcase the power of the nonparametric database-driven approach by constructing an accurate potential and using it to compute atomic-scale properties that are inaccessible to DFT due to computational expense. Second, while there has been vigorous activity recently in developing such models, most of the attention has been focused on the interpolation method and the neighborhood descriptors (e.g., neural networks [22–24], Shepherd interpolation [25,26], invariant polynomials [27–29], and Gaussian processes [30–34]); rather less prominence was given to the question of how to construct suitable databases that ultimately determine the range of validity of the potential. Our second goal is therefore to study what kinds of configurations need to be in a database so that given material properties are well reproduced. A larger database costs more to create and the resulting potential is slower, but can be expected to be more widely applicable, thus providing a tunable trade-off between transferability, accuracy, and computational cost.

In our Gaussian approximation potential (GAP) framework [30,31], the only uncontrolled approximation is the one essential to the idea of interatomic potentials: the total energy is written as a sum of atomic energies,

$$E = \sum_i \varepsilon(\hat{\mathbf{q}}_i), \tag{1}$$

with ε a universal function of the atomic neighborhood structure inside a finite cutoff radius as represented by the

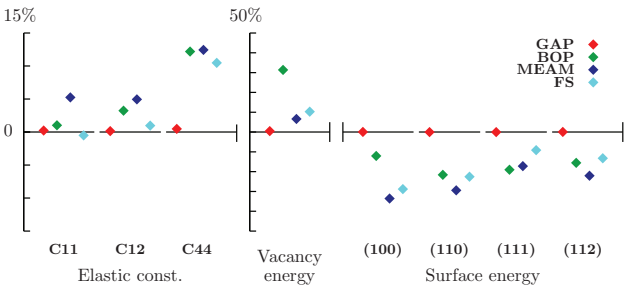

FIG. 1. (Color online) Fractional error in elastic constants and defect energies calculated with various interatomic potentials, as compared to the target DFT values.

[1]Rescaled to the DFT lattice constant and bulk modulus.

TABLE I. Elastic constants and defect energies calculated with various interatomic potentials, and corresponding target DFT values.

	DFT	GAP	BOP	MEAM	FS
C_{11} (GPa)	517	518	522	544	514
C_{12} (GPa)	198	198	205	208	200
C_{44} (GPa)	142	143	160	160	157
Vacancy energy (eV)	3.27	3.29	4.30	3.49	3.61
100 surface (eV/Å^2)	0.251	0.252	0.221	0.167	0.179
110 surface (eV/Å^2)	0.204	0.204	0.160	0.144	0.158
111 surface (eV/Å^2)	0.222	0.222	0.180	0.184	0.202
112 surface (eV/Å^2)	0.216	0.216	0.182	0.168	0.187

descriptor vector $\hat{\mathbf{q}}_i$ for atom i (defined below). This function is fitted to a database of DFT calculations using Gaussian process regression [35,36] so, in general, it is given by a linear combination of basis functions,

$$\varepsilon(\hat{\mathbf{q}}) = \sum_j \alpha_j K(\hat{\mathbf{q}}_j, \hat{\mathbf{q}}) \equiv \mathbf{k}(\hat{\mathbf{q}})^T \boldsymbol{\alpha}, \tag{2}$$

where the sum over j includes (some or all of) the configurations in the database, the vector of coefficients $\boldsymbol{\alpha}$ is given by linear algebra expressions (see below and in [30]), and the meaning of the covariance kernel K is that of a similarity measure between different neighbor environments.

The expression for the coefficients α_j—normally simple in Gaussian process regression—is more complicated in our case because the quantum mechanical input data we can calculate are not a set of values of the atomic energy function that we are trying to fit. Rather, the total energy of a configuration is a sum of many atomic energy function values, and the forces and stresses, which are also available analytically through the Hellmann-Feynman theorem, are sums of partial derivatives of the atomic energy function. The detailed derivation of the formulas shown below is in [38–40]. Let us collect all the input data values (total energies and force and stress components) into the vector $\mathbf{y}$ with D components in total and denote by $\mathbf{y}'$ the N *unknown* atomic energy values corresponding to all the atoms that appear in all the input configurations. We construct a linear operator $\mathbf{L}$ that describes the relationship between them through $\mathbf{y} = \mathbf{L}^T\mathbf{y}'$. For data values that represent total energies, the corresponding rows of $\mathbf{L}$ have just 0's and 1's as their elements, but for forces and stresses, the entries are differential operators such as $\partial/\partial x_i$ corresponding to the force on atom i with Cartesian x coordinate x_i. Writing $K_{ij} \equiv K(\hat{\mathbf{q}}_i, \hat{\mathbf{q}}_j)$ for the element of the covariance matrix $\mathbf{K}_{NN}$ corresponding to atoms i and j, the covariance matrix of size $D \times D$ of the observed data is

$$\mathbf{K}_{DD} = \mathbf{L}^T\mathbf{K}_{NN}\mathbf{L}, \tag{3}$$

where the differential operators in $\mathbf{L}$ act on the covariance function K that defines $\mathbf{K}_{NN}$. In our applications, N can exceed 100 000, and therefore working with $N \times N$ matrices would be computationally very expensive. Because many atomic environments in our data set are highly similar to one another, it is plausible that many fewer than N atoms could be chosen to efficiently represent the range of neighbor environments. We choose M representative atoms from the full set of N atoms that appear in all the input configurations (typically with $M \ll N$), and denote the square covariance matrix between the M representative atoms by $\mathbf{K}_{MM}$ and the rectangular covariance matrix between the M representative atoms and all the N atoms by $\mathbf{K}_{MN}$ (with $\mathbf{K}_{NM} = \mathbf{K}_{MN}^T$). The expression for the vector of coefficients in Eq. (2) is then

$$\boldsymbol{\alpha} = [\mathbf{K}_{MM} + \mathbf{K}_{MN}\mathbf{L}\Lambda^{-1}\mathbf{L}^T\mathbf{K}_{NM}]^{-1}\mathbf{K}_{MN}\mathbf{L}\Lambda^{-1}\mathbf{y}, \tag{4}$$

with

$$\Lambda = \sigma_\nu^2\mathbf{I}, \tag{5}$$

where the parameter σ_ν represents the tolerance (or expected error) in fitting the input data. It could be a single constant, but in practice we found it essential to use different tolerance values corresponding to the different kinds of input data, so that the Λ matrix is still diagonal, but has different values corresponding to total energies, forces, and stresses as they appear in the data vector $\mathbf{y}$. Although one might initially expect zero error in *ab initio* input data, this is not actually the case due to convergence parameters in the electronic structure calculation. A further source of error in the fit is the uncontrolled approximation of Eq. (1), i.e. writing the total energy as a sum of local atomic energies. The numerical values we use are shown in Table II. They are based on convergence tests of the DFT calculation carried out on example configurations.

TABLE II. DFT parameters used to generate training data and GAP model parameters.

DFT code	CASTEP [37] (version 6.01)
Exchange-correlation functional	PBE
Pseudopotential	Ultrasoft (valence $5s^25p^65d^46s^2$)
Plane-wave energy cutoff	600 eV
Maximum k-point spacing	0.015 Å^{-1}
Electronic smearing scheme	Gaussian
Smearing width	0.1 eV
Atomic environment kernel	SOAP
$f_{\text{cut}}(r) = \begin{cases} 1, & 0 < r \leqslant (r_{\text{cut}} - r_\Delta), \\ \frac{1}{2}\left(1 + \cos\left(\pi \frac{r - r_{\text{cut}} + r_\Delta}{r_\Delta}\right)\right), & (r_{\text{cut}} - r_\Delta) < r \leqslant r_{\text{cut}}, \\ 0, & r_{\text{cut}} < r. \end{cases}$	
$\phi_n(r) = \exp\left[-(r - r_{\text{cut}}n/n_{\max})^2/2\sigma_{\text{atom}}^2\right]$	
$S_{nn'} = \int_0^{r_{\text{cut}}} dr\, r^2\phi_n(r)\phi_{n'}(r), \quad \mathbf{S} = \mathbf{U}^T\mathbf{U}$	
$g_n(r) = \sum_{n'}(\mathbf{U}^{-1})_{nn'}\phi_{n'}(r)$	
r_{cut}	5.0 Å
r_Δ	1.0 Å
$\sigma_\nu^{(\text{energy})}$	0.0001 eV/atom
$\sigma_\nu^{(\text{force})}$	0.01 eV/Å
$\sigma_\nu^{(\text{virial})}$	0.01 eV/atom
σ_w	1.0 eV
σ_{atom}	0.5 Å
ξ	4
$n_{\max}$	14
$l_{\max}$	14
GAP software version	`df1c4d9`

We note the following remarks about the expression in (4). The quantum mechanically undefined and therefore unknown atomic energies for the input configurations $\mathbf{y}'$ do not appear. The number of components in the coefficient vector $\boldsymbol{\alpha}$ is M, so the sum in Eq. (2) is over the M representative configurations. The cost of calculating $\boldsymbol{\alpha}$ is dominated by operations which scale as $O(NM^2)$, so it can be significantly reduced by choosing M to be smaller and accepting a reduced accuracy of the fit. After the fit is made the coefficient vector $\boldsymbol{\alpha}$ stays fixed, and the evaluation of the potential is accomplished by the vector dot product in (2) with most of the work going towards computing the vector $\mathbf{k}$ for each new configuration, and thus scaling as $O(M)$. The M representative atoms can be chosen randomly, but we found it beneficial to employ the k-means clustering algorithm to choose the representative configurations.

We now turn to the specification of the kernel function. We use the "smooth overlap of atomic positions" (SOAP) kernel [31],

$$K_{ij} = \sigma_w^2 |\hat{\mathbf{q}}_i \cdot \hat{\mathbf{q}}_j|^{\xi}, \tag{6}$$

where the exponent ξ is a positive integer parameter whose role is to "sharpen" the selectivity of the similarity measure, and σ_w is an overall scale factor. Note that for the special choice of $\xi = 1$, the Gaussian process regression fit is equivalent to simple linear regression, and so the potential energy expression in (2) simplifies to $\varepsilon(\hat{\mathbf{q}}) = (\sigma_w^2 \sum_j \alpha_j \hat{\mathbf{q}}_j) \cdot \hat{\mathbf{q}}$, in which the term in parentheses can be precomputed once and for all. Unfortunately we found that such a linear fit significantly limits the attainable accuracy of the potential.

The elements of the descriptor vector $\hat{\mathbf{q}}$ are constructed as follows. The environment of the ith atom is characterized by the *atomic neighbourhood density*, which we define as

$$\begin{aligned} \rho_i(\mathbf{r}) &= \sum_j e^{-|\mathbf{r}-\mathbf{r}_{ij}|^2/2\sigma_{\text{atom}}^2} f_{\text{cut}}(|\mathbf{r}_{ij}|) \\ &= \sum_{\substack{n<n_{\max} \\ l<l_{\max} \\ |m|\leq l}} c_{nlm}^i g_n(|\mathbf{r}|) Y_{lm}(\hat{\mathbf{r}}), \end{aligned} \tag{7}$$

where $\mathbf{r}_{ij}$ are the vectors pointing to the neighboring atoms, σ_{atom} is a parameter corresponding to the "size" of the atoms, f_{cut} is a smooth cutoff function with compact support, and the expansion on the second line uses spherical harmonics and a set of orthonormal radial basis functions g_n with n, l, and m the usual integer indices. The elements of the descriptor vector $\hat{\mathbf{q}}$ are then

$$\mathbf{q}_i = \left\{ \sum_m \left(c_{nlm}^i\right)^* c_{n'lm}^i \right\}_{nn'l}, \quad \hat{\mathbf{q}}_i = \mathbf{q}_i / |\mathbf{q}_i|. \tag{8}$$

Values for the all the parameters and other necessary formulas are given in Table II. The orthonormal radial basis is obtained from a set of equispaced Gaussians by Cholesky factorization of their overlap matrix.

The SOAP kernel is special because it is not only invariant with respect to relabeling of atoms and rotation of either neighbor environment, but also faithful in the sense that K takes the value of unity only when the two neighborhoods are identical. This is because it is directly proportional to the overlap of the atomic neighborhood densities, integrated over all three-dimensional rotations $\hat{R}$ [31],

$$K_{ij} \propto \left| \int d\hat{R} \left| \int d\mathbf{r} \rho_i(\mathbf{r}) \rho_j(\hat{R}\mathbf{r}) \right|^2 \right|^{\xi}. \tag{9}$$

The SOAP kernel is therefore also manifestly smooth and slowly varying in Cartesian space, just as we know the true Born-Oppenheimer potential energy surface to be, away from electronic energy level crossings and quantum phase transitions. The entire GAP framework, including the choice of descriptor and the kernel, is designed so that its parameters are easy to set and the final potential is not very sensitive to the exact values. Some are physically motivated and stem from either the properties of the quantum mechanical potential energy surface (r_{cut}, σ_w, σ_{atom}) or the input data (e.g., σ_v), while others are convergence parameters and are set by a trade-off between accuracy and computational cost ($n_{\max}$, $l_{\max}$, M). We include in the Supplemental Information [41] a table demonstrating convergence of the fitted potential as a function of $n_{\max}$, $l_{\max}$, and r_{cut}. By far the most "arbitrary" part of the potential is thus the set of configurations chosen to comprise the training database.

Since the potential interpolates the atomic energy in the space of neighbor environments, we need good coverage of *relevant* environments in the database. We therefore need to start by deciding what material properties we wish to study and what are the corresponding neighbor environments. Our strategy is to define, for each material property, a set of representative *small unit cell* configurations that are amenable to accurate first-principles calculation. In Table III we show the performance with respect to key material properties of six models, each fitted to a database that contains the configurations indicated on the left, in addition to all the configurations of the preceding one. In particular, as proposed by Vitek and co-workers [42–44], the structure of $\frac{1}{2}\langle 111\rangle$ screw dislocations in bcc transition metals can be rationalized in terms of the strictly planar γ surface concept, and therefore we use γ surfaces in the database to ensure the coverage of neighbor environments found near the dislocation core. Where the dislocation structure is very far from correct, the numerical performance metric on it has been omitted. The table shows that, broadly speaking, the small representative unit cells are necessary and also sufficient to obtain each property accurately, so the GAP model interpolates well but does not extrapolate to completely new kinds of configurations. Adding new configurations never compromises the accuracy of previously incorporated properties. For information, Table IV shows the results of the automatic allocation of the representative atoms in each GAP model to the various types of configurations.

We also show the performance of the final GAP$_6$ model on Fig. 1 and omit the subscript from now. The phonon spectrum of the GAP model is shown in Fig. 2 along with that of the DFT and FS potential. There is clear improvement with respect to the analytical model, but remaining deficiencies are also apparent. Strategies to enhance the training database in order to improve the description of phonons are an important future direction of study.

TABLE III. (Color online) Summary of the databases for six GAP models, in order of increasing breadth in the types of configurations they contain, together with the performance of the corresponding potentials with respect to key properties. The color of the cells indicates a subjective judgment of performance: unacceptable (red), usable (yellow), good (green). The first five properties can be checked against DFT directly and so we report errors, but calculation of the last two properties are in large systems, so we report the values, converged with system size. The configurations are collected using Boltzmann sampling; for more details on the databases leading to the models see the Supplemental Information [41].

	Database	Computational cost[a] (ms/atom)	Elastic constants[b] (GPa)	Phonon spectrum[b] (THz)	Vacancy formation[c] (eV)	Surface energy[b] (eV/Å^2)	Dislocation structure[d] (Å^{-1})	Dislocation-vacancy binding energy (eV)	Peierls barrier (eV/b)
GAP_1 :	2000 × primitive unit cell with varying lattice vectors	24.70	0.623	*0.583*	*2.855*	*0.1452*	*0.0008*		
GAP_2 :	GAP_1 + 60 × 128-atom unit cell	51.05	0.608	0.146	*1.414*	*0.1522*	*0.0006*		
GAP_3 :	GAP_2 + vacancy in: 400 × 53-atom unit cell, 20 × 127-atom unit cell	63.65	0.716	0.142	0.018	*0.0941*	*0.0004*		
GAP_4 :	GAP_3 + (100), (110), (111), (112) surfaces 180 × 12-atom unit cell; (110), (112) γ surfaces 6183 × 12-atom unit cell	86.99	0.581	0.138	0.005	0.0001	0.0002	*−0.960*	*0.108*
GAP_5 :	GAP_4 + vacancy in: (110), (112) γ surface 750 × 47-atom unit cell	93.86	0.865	0.126	0.011	0.0001	0.0002	−0.774	*0.154*
GAP_6 :	GAP_5 + $\frac{1}{2}\langle 111\rangle$ dislocation quadrupole 100 × 135-atom unit cell	93.33	0.748	0.129	0.015	0.0001	0.0001	−0.794	0.112

[a]Time on a single CPU core of Intel Xeon E5-2670 2.6 GHz.
[b]rms error.
[c]Formation energy error.
[d]rms error of Nye tensor over the 12 atoms nearest the dislocation core; cf. Fig 4.

We now investigate the properties of the $\frac{1}{2}\langle 111\rangle$ screw dislocation further by calculating the Peierls barrier using a transition-state-searching implementation of the string method [45,46]. Three different initial transition paths, shown in Fig. 3, are used to explore the existence of the metastable state corresponding to a "hard" core structure [15,47–49]. We find that the hard core is not even locally stable in tungsten—starting geometry optimization from there results

TABLE IV. Number of representative atomic environments in each database of the six GAP models. The rows represent the successive GAP models and the columns represent the configuration types in the databases, grouped according to which GAP model first incorporated them. The allocations shown are based on *k*-means clustering. The rightmost column shows the total number of representative atoms in each GAP model (M).

	Database						Total
	1	2	3	4	5	6	M
GAP_1	2000						2000
GAP_2	814	3186					4000
GAP_3	366	1378	4256				6000
GAP_4	187	617	1890	6306			9000
GAP_5	158	492	1604	5331	2415		10000
GAP_6	140	450	1500	4874	2211	825	10000

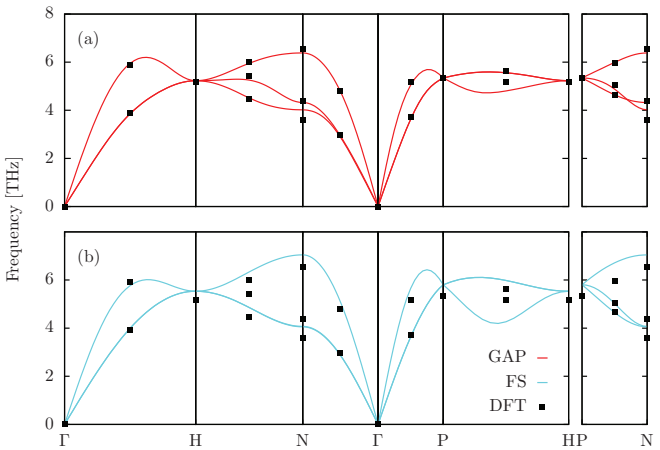

FIG. 2. (Color online) Phonon spectrum of bcc tungsten calculated using GAP and FS potentials, and some reference DFT values.

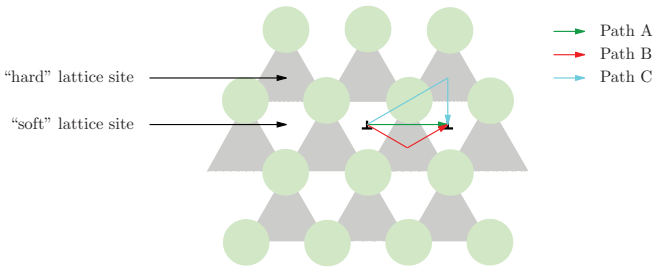

FIG. 3. (Color online) Representation of the three different initial transition paths for the Peierls barrier calculation. Path A corresponds to the linear interpolation directly from the initial to the final state, whereas paths B and C are the two distinct linear interpolations that include a potential metastable state (corresponding to the hard structure of the dislocation core) at reaction coordinate $r = 0.5$.

in the dislocation line migrating to a neighboring lattice site, corresponding to the "soft" core configuration. All three initial transition paths converge to the same minimum energy pathway (MEP), shown in Fig. 4, with no hard core transition state. For large enough systems, the MEP is independent of the boundary conditions: the "quadrupole" calculations contained two oppositely directed dislocations in periodic boundary conditions, while the "cylinder" configurations had a single dislocation with fixed far-field boundary conditions. For comparison we also plot the MEP of the Finnis-Sinclair model, and show the corresponding core structures using Nye tensor maps [50,51]. For the smallest periodic 135-atom model, we computed the energies at five points along the MEP using DFT to verify that the GAP model is indeed accurate for these configurations.

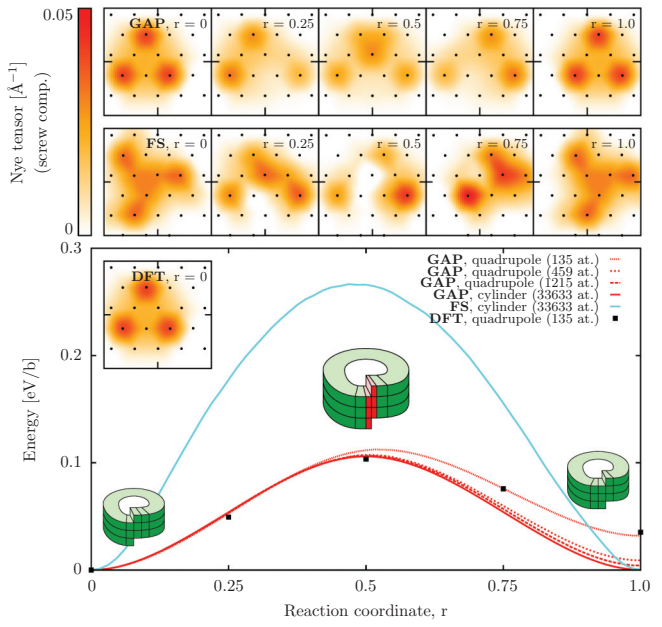

FIG. 4. (Color online) Top: The structure of the screw dislocation along the minimum energy path as it glides. Bottom: Peierls barrier evaluated using GAP and FS potentials, along with single-point checks with DFT in the 135-atom quadrupole arrangement.

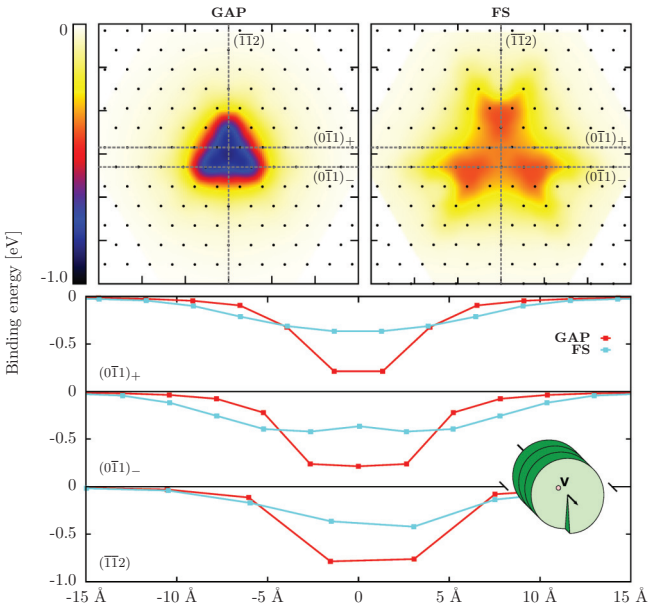

FIG. 5. (Color online) Dislocation-vacancy binding energy evaluated using GAP and FS potentials. The top panels show the interpolated binding energy using a heat map; the graphs below are slices of the same along the dotted lines shown in the top panels.

Due to the intrinsic smoothness of the potential, it can be expected to perform well for configurations which contain multiple defect structures as long as the local deformation around each defect with respect to the corresponding configurations in the database is small. So we finally turn to an example of the kinds of atomistic properties that are needed to make the connection to materials modeling on higher length scales, but are inaccessible to direct DFT calculations due to system size limitations imposed by the associated computational cost. Figure 5 shows the energy of a vacancy in the vicinity of a screw dislocation calculated in a system of over 100 000 atoms using cylindrical fixed boundary conditions 230 Å away from the core and with periodic boundary conditions applied along the dislocation line with a periodicity corresponding to three Burgers vectors. The Finnis-Sinclair potential underestimates this interaction by a factor of 2.

Although the potential developed in this work does not yet constitute a comprehensive description of tungsten under all conditions, we have shown that the strategy of building a database of representative small unit cell configurations is viable, and will be continued with the incorporation of other crystal phases, edge dislocations, interstitials, etc. In addition to developing ever more comprehensive databases and computing specific atomic scale properties with first-principles accuracy on which higher-length-scale models can be built, our long-term goal is to discover whether, in the context of a given material, an all-encompassing database could be assembled that contains a sufficient variety of neighbor environments to be valid for any configuration encountered under conditions of physically realistic temperatures and pressures. If that turns out to be possible, it will herald a truly new era of precision for atomistic simulations in materials science.

The authors are indebted to A. De Vita and N. Bernstein for comments on the manuscript. A.P.B. is supported by a Leverhulme Early Career Fellowship and the Isaac Newton Trust. G.C. acknowledges support from the EPSRC Grants No. EP/J010847/1 and No. EP/L014742/1. All software and data necessary for the reproduction of the results in this paper are available at www.libatoms.org.

[1] M. W. Finnis and J. E. Sinclair, Philos. Mag. A **50**, 45 (1984).
[2] M. S. Daw and M. I. Baskes, Phys. Rev. B **29**, 6443 (1984).
[3] G. J. Ackland and R. Thetford, Philos. Mag. A **56**, 15 (1987).
[4] A. P. Sutton and J. Chen, Philos. Mag. Lett. **61**, 139 (1990).
[5] J. Wang, Y. L. Zhou, M. Li, and Q. Hou, Modell. Simul. Mater. Sci. Eng. **22**, 015004 (2014).
[6] F. Ercolessi and J. B. Adams, Europhys. Lett. **26**, 583 (1994).
[7] M. I. Baskes, Phys. Rev. B **46**, 2727 (1992).
[8] Y. R. Wang and D. B. Boercker, J. Appl. Phys. **78**, 122 (1995).
[9] B.-J. Lee, M. I. Baskes, H. Kim, and Y. K. Cho, Phys. Rev. B **64**, 184102 (2001).
[10] M.-C. Marinica, L. Ventelon, M. R. Gilbert, L. Proville, S. L. Dudarev, J. Marian, G. Bencteux, and F. Willaime, J. Phys.: Condens. Matter **25**, 395502 (2013).
[11] M. Mrovec, R. Gröger, A. G. Bailey, D. Nguyen-Manh, C. Elsässer, and V. Vitek, Phys. Rev. B **75**, 104119 (2007).
[12] T. Ahlgren, K. Heinola, N. Juslin, and A. Kuronen, J. Appl. Phys. **107**, 033516 (2010).
[13] X.-C. Li, X. Shu, Y.-N. Liu, F. Gao, and G.-H. Lu, J. Nucl. Mater. **408**, 12 (2011).
[14] J. A. Moriarty, Phys. Rev. B **38**, 3199 (1988).
[15] W. Xu and J. A. Moriarty, Phys. Rev. B **54**, 6941 (1996).
[16] M. Mrovec, D. Nguyen-Manh, D. G. Pettifor, and V. Vitek, Phys. Rev. B **69**, 094115 (2004).
[17] G. F. Matthews, P. Edwards, T. Hirai, M. Kear, A. Lioure, P. Lomas, A. Loving, C. Lungu, H. Maier, P. Mertens *et al.*, Phys. Scr. **2007**, 137 (2007).
[18] R. Neu, M. Balden, V. Bobkov, R. Dux, O. Gruber, A. Herrmann, A. Kallenbach, M. Kaufmann, C. F. Maggi, H. Maier *et al.*, Plasma Phys. Controlled Fusion **49**, B59 (2007).
[19] R. Pitts, S. Carpentier, F. Escourbiac, T. Hirai, V. Komarov, S. Lisgo, A. S. Kukushkin, A. Loarte, M. Merola, A. Sashala *et al.*, J. Nucl. Mater. **438** (Suppl.), S48 (2013).
[20] N. Bernstein, J. R. Kermode, and G. Csányi, Rep. Prog. Phys. **72**, 026501 (2009).
[21] A. D. Vita and R. Car, MRS Bull. **491**, 473 (1997).
[22] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).
[23] J. Behler, R. Martoňák, D. Donadio, and M. Parrinello, Phys. Rev. Lett. **100**, 185501 (2008).
[24] N. Artrith and J. Behler, Phys. Rev. B **85**, 045439 (2012).
[25] J. Ischtwan and M. A. Collins, J. Chem. Phys. **100**, 8080 (1994).
[26] M. A. Collins, Theor. Chem. Acc. **108**, 313 (2002).
[27] X. Zhang, S. Zou, L. B. Harding, and J. M. Bowman, J. Phys. Chem. A **108**, 8980 (2004).
[28] X. Huang, B. J. Braams, and J. M. Bowman, J. Chem. Phys. **122**, 044308 (2005).
[29] Z. Xie, B. J. Braams, and J. M. Bowman, J. Chem. Phys. **122**, 224307 (2005).
[30] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Phys. Rev. Lett. **104**, 136403 (2010).
[31] A. P. Bartók, R. Kondor, and G. Csányi, Phys. Rev. B **87**, 184115 (2013).
[32] A. P. Bartók, M. J. Gillan, F. R. Manby, and G. Csányi, Phys. Rev. B **88**, 054104 (2013).
[33] M. J. Gillan, D. Alfè, A. P. Bartók, and G. Csányi, J. Chem. Phys. **139**, 244504 (2013).
[34] M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, Phys. Rev. Lett. **108**, 058301 (2012).
[35] D. MacKay, *Information Theory, Inference and Learning Algorithms* (Cambridge University Press, Cambridge, 2003).
[36] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning* (MIT Press, Cambridge, MA, 2006).
[37] S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson, and M. C. Payne, Z. Krystallogr. **220**, 567 (2005).
[38] E. Snelson and Z. Ghahramani, in *Advances in Neural Information Processing Systems* (MIT Press, Cambridge, MA, 2006), Vol. 18, pp. 1257–1264.
[39] A. P. Bartók, Ph.D. thesis, University of Cambridge, 2010.
[40] W. J. Szlachta, Ph.D. thesis, University of Cambridge, 2013.
[41] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.90.104108 for convergence of GAP_3 model as a function of parameters of the underlying covariance function.
[42] V. Vitek and F. Kroupa, Philos. Mag. **19**, 265 (1969).
[43] V. Vitek, R. C. Perrin, and D. K. Bowen, Philos. Mag. **21**, 1049 (1970).
[44] V. Vitek, Philos. Mag. **84**, 415 (2004).
[45] W. E, W. Ren, and E. Vanden-Eijnden, Phys. Rev. B **66**, 052301 (2002).
[46] Weinan E, W. Ren, and E. Vanden-Eijnden, J. Chem. Phys. **126**, 164103 (2007).
[47] S. Ismail-Beigi and T. A. Arias, Phys. Rev. Lett. **84**, 1499 (2000).
[48] D. E. Segall, A. Strachan, W. A. Goddard, S. Ismail-Beigi, and T. A. Arias, Phys. Rev. B **68**, 014104 (2003).
[49] D. Cereceda, A. Stukowski, M. R. Gilbert, S. Queyreau, L. Ventelon, M.-C. Marinica, J. M. Perlado, and J. Marian, J. Phys.: Condens. Matter **25**, 085702 (2013).
[50] C. Hartley and Y. Mishin, Acta Mater. **53**, 1313 (2005).
[51] B. G. Mendis, Y. Mishin, C. S. Hartley, and K. J. Hemker, Philos. Mag. **86**, 4607 (2006).