

To appear in *Philosophical Magazine*  
Vol. 00, No. 00, Month 20XX, 1–22

## Diffusivity and derivatives for interstitial solutes: Activation energy, volume, and elastodiffusion tensors

Dallas R. Trinkle\*

*Department of Materials Science and Engineering, University of Illinois, Urbana-Champaign*

(July 11, 2016)

Computational atomic-scale methods continue to provide new information about geometry, energetics, and transition states for interstitial elements in crystalline lattices. This data can be used to determine the diffusivity of interstitials by finding steady-state solutions to the master equation. In addition, atomic-scale computations can provide not just the site energy, but also the stress in the cell due to the introduction of the defect to compute the elastic dipole. We derive a general expression for the fully anisotropic diffusivity tensor from site and transition state energies, and three derivatives of the diffusivity: the elastodiffusion tensor (derivative of diffusivity with respect to strain), the activation barrier tensor (logarithmic derivative of diffusivity with respect to inverse temperature) and activation volume tensor (logarithmic derivative of diffusivity with respect to pressure). Computation of these quantities takes advantage of crystalline symmetry, and we provide an open-source implementation of the algorithm. We provide analytic results for octahedral-tetrahedral networks in face-centered cubic, body-centered cubic, and hexagonal closed-packed lattices, and conclude with numerical results for C in Fe.

**Keywords:** Interstitial diffusion; activation barrier; elastodiffusion tensor; automated computation

### 1. Introduction

Mass transport in solids is an integral part of material processing for materials from metals to ceramics to semiconductors. At the atomic scale, atoms move via defects: either vacancies for atoms on the crystalline lattice, or via interstitial sites off of the lattice[1, 2]. For the case of interstitial defects, what had been often considered the “simplest” transport to model can often conceal surprising complexity[3] when the interstitial network involves more than a single symmetry unique site. Furthermore, real materials also have non-homogeneous strain fields, and strain provides both a driving force[4] for diffusion and modifies transport coefficients leading to a variety of complex behavior including: anisotropic dopant diffusion in semiconductor thin films[5–7], internal stress fields near a dislocation affecting transport and segregation of C solute atoms in Fe[8, 9] or vacancies in HCP metals[10], and anisotropy of dumbbell interstitial diffusion under biaxial stress in Cu and Pt[11]. The influence of strain on diffusivity has been previously investigated only for cases where the interstitial random walk is purely uncorrelated[12, 13]; this excludes even the simple case of octahedral-tetrahedral networks in hexagonal-closed packed materials. While advances in density functional theory using modern supercomputers combined with transition-state finding methods[14, 15] can compute the energies for different con-

---

\* Email: dtrinkle@illinois.edu

figurations and transition rates at the atomic scale with near chemical accuracy[16, 17], general approaches to (a) automate the generation of a transition network, and (b) compute transport coefficients and related quantities using non-stochastic techniques are lacking.

The simplest transport coefficient to consider is the (generally anisotropic) diffusivity tensor  $\underline{D}$ , but we can also conceive of derivatives with respect to temperature and strain (or stress). An interstitial solute executes a random walk over a connected network of sites, where the transition rate between sites is given by a rate following an Arrhenius relation from transition-state theory[18]. From the long-time behavior, the diffusivity tensor  $\underline{D}$  can be computed, where  $\underline{D}$  relates a steady-state flux  $\mathbf{j}$  to a concentration gradient  $-\nabla c$ . The existence of different symmetry unrelated sites (and transitions) means that the total diffusivity may not have a simple Arrhenius temperature dependence given by a single activation barrier. Instead, over a sufficiently small temperature range, if we were to approximate the diffusivity tensor with an Arrhenius relation,  $\underline{D} \approx \underline{D}_0 \exp(-\beta \underline{E}^{\text{act}})$ , then the negative derivative of diffusivity with respect to inverse temperature  $\beta$  allows us to define

$$\underline{E}^{\text{act}} := -\underline{D}^{-1/2} \frac{d\underline{D}}{d\beta} \underline{D}^{-1/2}, \quad (1)$$

corresponding to an anisotropic (and possibly temperature dependent) “barrier” tensor. The barrier should correspond to the activation energy for the rate-limiting transition in diffusion. The fourth-rank elastodiffusion tensor[1, 13],

$$\underline{d} := \frac{d\underline{D}}{d\boldsymbol{\varepsilon}}, \quad (2)$$

is the derivative of diffusivity with respect to strain. As both diffusivity and strain are symmetric tensors,  $d_{abcd} = d_{bacd} = d_{abdc}$ ; furthermore,  $\underline{d}$  obeys crystalline symmetry. Thus, it has a similar structure to elastic stiffnesses and compliances, with one exception: as it is not the second derivative of an energy, there is no general relationship between  $d_{abcd} = dD_{ab}/d\varepsilon_{cd}$  and  $d_{cdab} = dD_{cd}/d\varepsilon_{ab}$ , unless imposed by crystalline symmetry. Finally, activation volume can be defined in a similar fashion to activation energy, as a logarithmic derivative with respect to stress,

$$\underline{V}_{abcd}^{\text{act}} := k_B T \underline{D}^{-1/2} \frac{d\underline{D}}{d\boldsymbol{\sigma}} \underline{D}^{-1/2} = k_B T \sum_{ijkl=1}^3 \left( \underline{D}^{-1/2} \right)_{ai} \frac{d\underline{D}_{ij}}{d\varepsilon_{kl}} \left( \underline{D}^{-1/2} \right)_{jb} S_{klcd}, \quad (3)$$

and related back to the elastodiffusion tensor using the elastic compliances  $S_{ijkl}$ . This fourth-rank tensor contains the full anisotropy of diffusion response to stress. It can be reduced to a scalar activation volume by a double contraction,

$$V^{\text{act}} := \frac{1}{3} \sum_{ij=1}^3 \underline{V}_{iijj}^{\text{act}} \quad (4)$$

where the second contraction considers hydrostatic pressure, and the first contraction averages in all spatial directions.

In what follows, we develop the theory for the diffusivity, elastodiffusivity, activation barrier and volume tensors of an interstitial in terms of the site energies, transition state energies, and elastic dipoles (derivative of energy with respect to strain) of those same

quantities. This includes the cases where the interstitial random walk includes correlation. We apply the theory to three simple examples: octahedral-tetrahedral networks in face-centered cubic, body-centered cubic, and hexagonal closed-packed lattices. We also provide numerical values for carbon diffusion in iron, using data from[9]. Finally, the work outlined here for the general case is implemented in open-source, publicly-available code (c.f. Appendix C and [19]) that can be applied to other systems[20].

## 2. Methodology

### 2.1. Diffusivity

In the case of interstitial diffusivity in the dilute limit, an interstitial atom moves through a network of sites in the crystal. Such a model allows simplification of the master equation: the system state is full described by the location of the interstitial,  $\mathbf{x} = \mathbf{R} + \mathbf{u}$  for  $\mathbf{R}$  a lattice vector and  $\mathbf{u}$  a vector in the unit cell. Due to translational invariance, the equilibrium site probability and set of transitions from a site are determined by the vector in the unit cell. For a site  $i$  in the unit cell, it has an equilibrium site probability  $\rho_i$  that follows an Arrhenius relationship,

$$\rho_i := Z^{-1} \rho_i^0 \exp(-\beta E_i) \quad (5)$$

for site energy  $E_i$ , entropic prefactor  $\rho_i^0 = \exp(S_i/k_B)$ , and partition function  $Z = \sum_i \rho_i^0 \exp(-\beta E_i)$ . The transition from site  $i$  to site  $j$  has a rate  $\lambda_{i \rightarrow j}$ ,

$$\lambda_{i \rightarrow j} := \frac{\lambda_{ij}^0}{\rho_i^0} \exp(-\beta [E_{ij}^{\text{ts}} - E_i]) \quad (6)$$

for transition state energy  $E_{ij}^{\text{ts}}$  and entropic prefactor  $\lambda_{ij}^0 = \exp(S_{ij}^{\text{ts}}/k_B)$ , following [18]. In this formulation, the transition state energy and entropic prefactors are equal for  $i \rightarrow j$  and for  $j \rightarrow i$ , while it is not necessary that  $\lambda_{i \rightarrow j}$  and  $\lambda_{j \rightarrow i}$  are equal. Finally, the probabilities obey detailed balance, where  $\rho_i \lambda_{i \rightarrow j} = \rho_j \lambda_{j \rightarrow i}$  for all  $i, j$ . Each transition results in a displacement of the diffusing atom by

$$\delta \mathbf{x}_{i \rightarrow j} := \mathbf{x}_j - \mathbf{x}_i. \quad (7)$$

With these definitions of site probabilities and rates, we can write down an expression for the diffusivity[2]. There are two contributions to the diffusivity: an uncorrelated diffusivity and a correlation correction due to the unbalanced hops from individual sites. We define two rate matrices,  $\Lambda_{ij}$ ,

$$\Lambda_{ij} := \begin{cases} \lambda_{i \rightarrow j} & : i \neq j \\ -\sum_j \lambda_{i \rightarrow j} & : i = j \end{cases} \quad (8)$$

which is the transition matrix for the master equation, and its symmetric counterpart

$$\omega_{ij} := \rho_i^{1/2} \Lambda_{ij} \rho_j^{-1/2}. \quad (9)$$

These matrices are negative-definite, and—due to detailed balance—both have a null-space related to the site probabilities  $\rho_i$ . We define the scaled velocity vector (corresponding to

the bias of jumps at a site as in [8, 9]),

$$\mathbf{b}_i := \rho_i^{1/2} \sum_j \lambda_{i \rightarrow j} \delta \mathbf{x}_{i \rightarrow j} \quad (10)$$

which is non-zero when there are unbalanced hops from a site  $i$ ; together with the bias-correction vector  $\boldsymbol{\gamma}_i$  which solves

$$\sum_j \omega_{ij} \boldsymbol{\gamma}_j = \mathbf{b}_i. \quad (11)$$

This is most easily done with the pseudoinverse of  $\omega$ ,  $g$ , so that  $\boldsymbol{\gamma}_i = \sum_j g_{ij} \mathbf{b}_j$ . Finally, the diffusivity is

$$\underline{D} = \frac{1}{2} \sum_{ij} \delta \mathbf{x}_{i \rightarrow j} \otimes \delta \mathbf{x}_{i \rightarrow j} \lambda_{i \rightarrow j} \rho_i + \sum_i \mathbf{b}_i \otimes \boldsymbol{\gamma}_i \quad (12)$$

where  $\otimes$  is the outer (or dyad) product of two vectors.<sup>1</sup> Because  $\omega$  is a symmetric matrix, so is  $g$ , and hence  $\underline{D}$  is a symmetric second-rank tensor.

While the expressions leading up to Eqn. 12 are general for non-interacting atoms hopping between sites, crystal symmetry can reduce the complexity of the expressions to be evaluated. We use the Seitz notation[21] for a symmetry operation  $\{R, \mathbf{t}\}$ , where for a point  $\mathbf{x}$ ,  $\{R, \mathbf{t}\}\mathbf{x} := R\mathbf{x} + \mathbf{t}$ . Then, the inverse  $\{R, \mathbf{t}\}^{-1} = \{R^{-1}, -R^{-1}\mathbf{t}\}$ . The full set of operations make up the *space group*, and at any site  $i$  there will be the subgroup of operations that leave that site fixed, its *point group*. If every site  $i$  has a position  $\mathbf{x}_i$ , then each group operation can be expressed as a matrix

$$\{R, \mathbf{t}\}_{ij} := \delta(\mathbf{x}_i - \{R, \mathbf{t}\}\mathbf{x}_j) \quad (13)$$

where  $\delta$  is the Kronecker delta. This matrix then applies the group operation  $\{R, \mathbf{t}\}$  to a site scalar  $f_i$  as  $\sum_j \{R, \mathbf{t}\}_{ij} f_j$ . A related matrix can be defined corresponding to vector at each site (such as  $\mathbf{b}_i$  and  $\boldsymbol{\gamma}_i$ ), where with the orthonormal basis vectors  $\{\mathbf{e}_a\}$ , then the matrix

$$\{R, \mathbf{t}\}_{ia,jb} := \{R, \mathbf{t}\}_{ij} (\mathbf{e}_a \cdot R \cdot \mathbf{e}_b) \quad (14)$$

transforms any site vector  $\mathbf{f}_i$  as  $\sum_a \mathbf{e}_a \sum_{jb} \{R, \mathbf{t}\}_{ia,jb} (\mathbf{f}_i \cdot \mathbf{e}_b)$ . All of our site scalars ( $\rho_i$ ), site vectors ( $\mathbf{b}_i$  and  $\boldsymbol{\gamma}_i$ ), and site-to-site matrices ( $\Lambda_{ij}$ ,  $\omega_{ij}$ , and  $g_{ij}$ ) share the crystal symmetry,

---

<sup>1</sup>The construction  $\mathbf{a} \otimes \mathbf{b}$  is a second rank tensor such that  $(\mathbf{a} \otimes \mathbf{b}) \cdot \mathbf{v} = (\mathbf{b} \cdot \mathbf{v})\mathbf{a}$  for any vector  $\mathbf{v}$ .

and so the following symmetry relations are true for any space group operation  $\{R, \mathbf{t}\}$ ,

$$\begin{aligned}
 \rho_i &= \sum_j \{R, \mathbf{t}\}_{ij} \rho_j, \\
 \mathbf{b}_i &= \sum_a \mathbf{e}_a \sum_{jb} \{R, \mathbf{t}\}_{ia,jb} (\mathbf{b}_j \cdot \mathbf{e}_b), \\
 \boldsymbol{\gamma}_i &= \sum_a \mathbf{e}_a \sum_{jb} \{R, \mathbf{t}\}_{ia,jb} (\boldsymbol{\gamma}_j \cdot \mathbf{e}_b), \\
 \Lambda_{ij} &= \sum_{kl} \{R, \mathbf{t}\}_{ik} \Lambda_{kl} \{R, \mathbf{t}\}_{lj}, \\
 \omega_{ij} &= \sum_{kl} \{R, \mathbf{t}\}_{ik} \omega_{kl} \{R, \mathbf{t}\}_{lj}, \\
 g_{ij} &= \sum_{kl} \{R, \mathbf{t}\}_{ik} g_{kl} \{R, \mathbf{t}\}_{lj}.
 \end{aligned} \tag{15}$$

The implication of these symmetries is that  $\rho_i$ ,  $\mathbf{b}_i$ , and  $\boldsymbol{\gamma}_i$  can be written as linear combinations of vectors left unchanged by *all*  $\{R, \mathbf{t}\}$  matrices, and that our matrices  $\Lambda$ ,  $\omega$ , and  $g$  can be entirely expanded in that same basis.

The basis functions for site scalars and site vectors come from the Wyckoff positions corresponding to sites in the network. Each Wyckoff position represents a full set of symmetry-related sites; any site scalar that obeys crystal symmetry will have the same value for each site corresponding to the same Wyckoff position. Hence, if we have a Wyckoff position that has  $N_W$  sites in the unit cell, its basis vector components will be  $N_W^{-1/2}$  for each site corresponding to that Wyckoff position, and 0 otherwise; there will be one basis function for each unique Wyckoff site *regardless of how many sites that may represent in the unit cell*. The Wyckoff positions also serve to construct a basis for site vectors combined with the point group operations for that site. Considering the point group operations that are available in a crystal, there are ten types: identity, a 2-, 3-, 4-, or 6-fold axis, mirror through a plane, or a 2-, 3-, 4-, or 6-fold axis combined with a mirror through that same plane. Note that a 2-fold axis combined with a mirror operation is inversion. As we are interested in vectors that remain *unchanged* by all of the point group operations for a site, we can easily generate the basis corresponding to each operation. For identity, we have any 3-dimensional basis; for a mirror operation, a 2-dimensional basis that spans the mirror plane; for a 2-, 3-, 4-, or 6-fold axis, a 1-dimensional basis corresponding to the axis; and for a 2-, 3-, 4-, or 6-fold axis combined with a mirror, there are no vectors left unchanged. We construct the vector space left unchanged by each operation, and intersect to generate the final vector space for a single site. The spanning vectors for that site can then be rotated to the other sites corresponding to the same Wyckoff position, and normalized in a similar way to the site scalar basis.

The crystal symmetry divides diffusion networks into three types based on the crystal symmetry: networks without correlation, networks with correlation and inversion, and networks with correlation but without inversion. As it is possible for the site vector basis to be an empty set (e.g., if each Wyckoff position point group possesses a 2-, 3-, 4-, or 6-fold axis combined with a mirror operation), this excludes correlation *explicitly by symmetry*. Next, there may be a non-empty site vector basis while the crystal has inversion. In this case, the transition matrices  $\Lambda$  or  $\omega$  expressed in the site vector basis are not singular. This is because the null-space vector of  $\Lambda$ —which corresponds to  $\rho_i$ —has zero projection into any basis vector. Finally, in the most general case, the transition matrices remain singular,

and so the construction of  $g$  requires the pseudoinverse of  $\omega$ .

## 2.2. Activation barrier

We define the activation energy tensor from the first derivative of the diffusivity with respect to inverse temperature, and evaluate with perturbation theory. In this case, we introduce a small change in inverse temperature,  $-\delta\beta$ , to  $\rho_i$  and  $\lambda_{i \rightarrow j}$ , then evaluate the first order change in  $\underline{D}$ . Then, the first order change in  $\rho_i$ ,  $\delta\rho_i$ , is given by

$$\begin{aligned}\rho_i + \delta\rho_i &= \frac{\rho_i e^{\delta\beta E_i}}{\sum_j \rho_j e^{\delta\beta E_j}} \\ &= \rho_i + \rho_i (E_i - \langle E \rangle) \delta\beta + O(\delta\beta^2)\end{aligned}\tag{16}$$

where the average energy  $\langle E \rangle = \sum_i \rho_i E_i$ . Next, the first order change in  $\lambda_{i \rightarrow j}$ ,  $\delta\lambda_{i \rightarrow j}$ , is given by

$$\begin{aligned}\lambda_{i \rightarrow j} + \delta\lambda_{i \rightarrow j} &= \lambda_{i \rightarrow j} e^{\delta\beta (E_{ij}^{\text{ts}} - E_i)} \\ &= \lambda_{i \rightarrow j} + \lambda_{i \rightarrow j} (E_{ij}^{\text{ts}} - E_i) \delta\beta + O(\delta\beta^2).\end{aligned}\tag{17}$$

From these, we can compute the related terms:  $\delta\mathbf{b}_i$  and  $\delta\omega$ . First,

$$\begin{aligned}\mathbf{b}_i + \delta\mathbf{b}_i &= (\rho_i + \delta\rho_i)^{1/2} \sum_j (\lambda_{i \rightarrow j} + \delta\lambda_{i \rightarrow j}) \delta\mathbf{x}_{i \rightarrow j} \\ &= \mathbf{b}_i + \rho_i^{1/2} \sum_j \lambda_{i \rightarrow j} \left[ E_{ij}^{\text{ts}} - \frac{1}{2} (E_i + \langle E \rangle) \right] \delta\mathbf{x}_{i \rightarrow j} \delta\beta + O(\delta\beta^2).\end{aligned}\tag{18}$$

Next for  $\delta\omega_{ii}$  we have

$$\begin{aligned}\delta\omega_{ii} &= - \sum_j \delta\lambda_{i \rightarrow j} \\ &= - \sum_j \lambda_{i \rightarrow j} (E_{ij}^{\text{ts}} - E_i) \delta\beta + O(\delta\beta^2)\end{aligned}\tag{19}$$

and for  $\delta\omega_{ij}$  ( $i \neq j$ ) we have

$$\begin{aligned}\delta\omega_{ij} &= (\rho_i + \delta\rho_i)^{1/2} (\lambda_{i \rightarrow j} + \delta\lambda_{i \rightarrow j}) (\rho_j + \delta\rho_j)^{-1/2} \\ &= \omega_{ij} \left[ E_{ij}^{\text{ts}} - \frac{1}{2} (E_i + E_j) \right] \delta\beta + O(\delta\beta^2).\end{aligned}\tag{20}$$

Finally, the first-order change in the pseudoinverse  $g$ ,  $\delta g$  is given by (c.f. Appendix A)

$$\begin{aligned}g + \delta g &= (\omega + \delta\omega)^{-1} \\ &= g - g \delta\omega g + O(\delta\beta^2).\end{aligned}\tag{21}$$

With all of these terms, we can compute the first-order correction to the diffusivity, which gives the derivative of  $\underline{D}$  with respect to  $-\beta$  by substituting into Eqn. 12,

$$\begin{aligned} -\frac{d\underline{D}}{d\beta} = & \frac{1}{2} \sum_{ij} \delta \mathbf{x}_{i \rightarrow j} \otimes \delta \mathbf{x}_{i \rightarrow j} \left[ E_{ij}^{\text{ts}} - \langle E \rangle \right] \lambda_{i \rightarrow j} \rho_i \\ & + \sum_i \left( \frac{\delta \mathbf{b}_i}{\delta \beta} \otimes \boldsymbol{\gamma}_i + \boldsymbol{\gamma}_i \otimes \frac{\delta \mathbf{b}_i}{\delta \beta} \right) - \sum_{ij} \boldsymbol{\gamma}_i \otimes \frac{\delta \omega_{ij}}{\delta \beta} \boldsymbol{\gamma}_j \end{aligned} \quad (22)$$

where the first term is the uncorrelated contribution and the remaining terms come from correlation. It should be noted that all of the perturbed site scalar, site vector, and matrices have the same symmetry as their unperturbed versions. A similarly structured solution will follow for the evaluation of the elastodiffusion tensor.

### 2.3. Elastodiffusion tensor

To evaluate the elastodiffusion tensor, we introduce a perturbation through strain which changes site and transition state energies, as defined by the elastic dipole tensor. The elastic dipole tensor  $\underline{P}_i$  for a site  $i$  is

$$\underline{P}_i := -\frac{dE_i}{d\underline{\varepsilon}}. \quad (23)$$

The elastic dipole can be conveniently evaluated in a supercell calculation from the stress in the cell: an interstitial is added to an initially undefected, unstressed supercell containing  $N$  atoms (with equilibrium volume  $V_0$  per atom), resulting in a stress  $\underline{\sigma}$ , then to first order in  $N^{-1}$ ,

$$\underline{P} \approx NV_0 \underline{\sigma}, \quad (24)$$

which is straightforward to evaluate with density-functional theory methods; e.g., see [22–25]. Similarly, the energy of a transition state can also change with strain, as dictated by the elastic dipole tensor for the transition state  $\underline{P}_{ij}^{\text{ts}}$  for the transition state between  $i$  and  $j$ ,

$$\underline{P}_{ij}^{\text{ts}} := -\frac{dE_{ij}^{\text{ts}}}{d\underline{\varepsilon}}. \quad (25)$$

This, too, can be approximated by the stress at the transition state in a supercell calculation as in Eqn. 24; e.g., see [9, 24]. The definitions of elastic dipoles allow the introduction of a small strain perturbation  $\delta \underline{\varepsilon}$  to produce site energies changes  $\delta E_i$  and transition energies  $\delta E_{ij}^{\text{ts}}$  as

$$\begin{aligned} \delta E_i &= -\underline{P}_i : \delta \underline{\varepsilon} \\ \delta E_{ij}^{\text{ts}} &= -\underline{P}_{ij}^{\text{ts}} : \delta \underline{\varepsilon} \end{aligned} \quad (26)$$

which is correct to first order in strain.<sup>2</sup>

---

<sup>2</sup>The double contraction sums over both indices of the two second-rank tensors:  $A : B = \sum_{ab} A_{ab} B_{ba}$ .

We define the elastodiffusion tensor from the first derivative of the diffusivity with respect to strain, and evaluate with perturbation theory. Strain produces changes in site energies and transition state energies as laid out above, but also modifies all vectors in the network. The purely geometric contribution is evaluated by expanding the diffusivity tensor out in components as  $\underline{D} = \sum_{ab} D_{ab} \mathbf{e}_a \otimes \mathbf{e}_b$  for an orthonormal basis  $\{\mathbf{e}_a\}$ ; then a (symmetric) strain perturbs the basis vectors by  $\delta \mathbf{e}_a = \sum_c \delta \varepsilon_{ac} \mathbf{e}_c$ , so that

$$\delta \underline{D}^{(\text{geom})} := \sum_{ab} \mathbf{e}_a \otimes \mathbf{e}_b \left( \sum_c D_{ac} \delta \varepsilon_{cb} + D_{bc} \delta \varepsilon_{ca} \right) \quad (27)$$

which produces a (symmetrized) contribution to the elastodiffusion tensor

$$d_{abcd}^{(\text{geom})} := \frac{1}{2} (\delta_{ad} D_{bc} + \delta_{ac} D_{bd} + \delta_{bc} D_{ad} + \delta_{bd} D_{ac}). \quad (28)$$

To get the uncorrelated and correlated changes that we add to this, we get the first order correction to  $\rho_i$ ,  $\delta \rho_i$  as

$$\begin{aligned} \rho_i + \delta \rho_i &= \frac{\rho_i e^{-\beta \delta E_i}}{\sum_j \rho_j e^{-\beta \delta E_j}} \\ &= \rho_i + \rho_i \beta (\underline{P}_i - \langle \underline{P} \rangle) : \delta \underline{\varepsilon} + O(\delta \underline{\varepsilon}^2) \end{aligned} \quad (29)$$

where the average elastic dipole  $\langle \underline{P} \rangle = \sum_i \rho_i \underline{P}_i$ . Next, the first order change in  $\lambda_{i \rightarrow j}$ ,  $\delta \lambda_{i \rightarrow j}$ , is given by

$$\begin{aligned} \lambda_{i \rightarrow j} + \delta \lambda_{i \rightarrow j} &= \lambda_{i \rightarrow j} e^{-\beta (\delta E_{ij}^{\text{ts}} - \delta E_i)} \\ &= \lambda_{i \rightarrow j} + \lambda_{i \rightarrow j} \beta (\underline{P}_{ij}^{\text{ts}} - \underline{P}_i) : \delta \underline{\varepsilon} + O(\delta \underline{\varepsilon}^2). \end{aligned} \quad (30)$$

From these, we can compute the related terms:  $\delta \mathbf{b}_i$ —ignoring contributions of the strain to the hop vectors—and  $\delta \omega$ . First,

$$\begin{aligned} \mathbf{b}_i + \delta \mathbf{b}_i &= (\rho_i + \delta \rho_i)^{1/2} \sum_j (\lambda_{i \rightarrow j} + \delta \lambda_{i \rightarrow j}) \delta \mathbf{x}_{i \rightarrow j} \\ &= \mathbf{b}_i + \rho_i^{1/2} \sum_j \delta \mathbf{x}_{i \rightarrow j} \lambda_{i \rightarrow j} \beta \left[ \underline{P}_{ij}^{\text{ts}} - \frac{1}{2} (\underline{P}_i + \langle \underline{P} \rangle) \right] : \delta \underline{\varepsilon} + O(\delta \underline{\varepsilon}^2). \end{aligned} \quad (31)$$

Next for  $\delta \omega_{ii}$  we have

$$\begin{aligned} \delta \omega_{ii} &= - \sum_j \delta \lambda_{i \rightarrow j} \\ &= - \sum_j \lambda_{i \rightarrow j} \beta (\underline{P}_{ij}^{\text{ts}} - \underline{P}_i) : \delta \underline{\varepsilon} + O(\delta \underline{\varepsilon}^2) \end{aligned} \quad (32)$$

and for  $\delta \omega_{ij}$  ( $i \neq j$ ) we have

$$\begin{aligned} \delta \omega_{ij} &= (\rho_i + \delta \rho_i)^{1/2} (\lambda_{i \rightarrow j} + \delta \lambda_{i \rightarrow j}) (\rho_j + \delta \rho_j)^{-1/2} \\ &= \omega_{ij} \beta \left[ \underline{P}_{ij}^{\text{ts}} - \frac{1}{2} (\underline{P}_i + \underline{P}_j) \right] : \delta \underline{\varepsilon} + O(\delta \underline{\varepsilon}^2). \end{aligned} \quad (33)$$



The expression for the first-order change in the pseudoinverse is the same as Eqn. 21; hence, we can combine all of the contributions into Eqn. 12 (similar to Eqn. 22) plus Eqn. 28 to get

$$\begin{aligned}
 d_{abcd} = & \frac{1}{2} (\delta_{ad} D_{bc} + \delta_{ac} D_{bd} + \delta_{bc} D_{ad} + \delta_{bd} D_{ac}) \\
 & + \frac{1}{2} \sum_{ij} \delta \mathbf{x}_{i \rightarrow j,a} \delta \mathbf{x}_{i \rightarrow j,b} \beta (P_{ij,cd}^{\text{ts}} - \langle P \rangle_{cd}) \lambda_{i \rightarrow j} \rho_i \\
 & + \sum_i \left( \frac{\delta \mathbf{b}_{i,a}}{\delta \varepsilon_{cd}} \gamma_{i,b} + \gamma_{i,a} \frac{\delta \mathbf{b}_{i,b}}{\delta \varepsilon_{cd}} \right) - \sum_{ij} \gamma_{i,a} \frac{\delta \omega_{ij}}{\delta \varepsilon_{cd}} \gamma_{j,b}
 \end{aligned} \tag{34}$$

where the first term is the geometric contribution, the second term is the uncorrelated contribution, and the remaining terms come from correlation.

Introducing a finite strain into the lattice can reduce the symmetry of the lattice, and couples with the anisotropy of the elastic dipole tensors. The dipole tensors can be expanded in a site *tensor* basis, similar to the site vector basis, using the point group for a single representative site of each Wyckoff position. We can also make a similar expansion for the transition states, by working with the “double point group”: the set of symmetry operations that leave *both* the initial and final positions unchanged. Despite the breaking of symmetry, we can see from Eqn. 34, that if a site  $i$  has an empty vector basis (so that both  $\mathbf{b}_i$  and  $\gamma_i$  are zero by symmetry), the correlation contributions to the elastodiffusion tensor are *also* zero. This means that the introduction of strain *does not require* the computation of  $\delta \mathbf{b}_i$  or  $\delta \omega_{ij}$  for any sites where  $\mathbf{b}_i$  and  $\omega_{ij}$  was not already computed. However, it *is* possible that particular values of  $\delta \varepsilon$  can produce  $\delta \mathbf{b}_i$  with *components* that would be zero by symmetry for  $\mathbf{b}_i$ ; the symmetry of  $\delta \mathbf{b}_i / \delta \varepsilon_{jk}$  is that of a third-rank tensor at site  $i$ .

### 3. Results

We can apply these results to three classic systems: octahedral-tetrahedral networks in face-centered cubic (FCC), body-centered cubic (BCC), and hexagonal closed-packed (HCP) structures. In the case of FCC and BCC, the network has zero site vectors basis (no correlation contribution), while the HCP network has a site vector basis oriented along the  $c$ -axis. In the case of BCC, tetrahedrals can be transition states, or states on their own. These three cases move from one, to two, to three different transition state energies. Finally, we conclude with a short set of numerical results for carbon in iron, using the EAM data generated by [9].

#### 3.1. Face-centered cubic

In a face-centered cubic lattice, the common interstitial sites are octahedral sites that can diffuse to tetrahedral sites, and vice versa. The face-centered cubic space group is  $Fm\bar{3}m$ , where the solvent atoms occupy  $4a$  Wyckoff positions, with octahedrals at  $4b$  and tetrahedrals at  $8c$ . [26] There is one octahedral site for every “solvent” atom, forming an interpenetrating FCC lattice. Each octahedral is connected to eight tetrahedral sites via  $\frac{a}{4}\langle 111 \rangle$  jumps. There are two tetrahedral sites for every lattice atom, forming an interpenetrating simple cubic lattice. Each tetrahedral is connected to four octahedral sites; the jump vectors are a subset of  $\frac{a}{4}\langle 111 \rangle$  vectors with tetrahedral symmetry (either  $\{[111], [1\bar{1}\bar{1}], [\bar{1}1\bar{1}], [\bar{1}\bar{1}1]\}$  or  $\{[\bar{1}\bar{1}\bar{1}], [\bar{1}11], [1\bar{1}1], [11\bar{1}]\}$ ). By symmetry, there is a single transition state energy,  $E_{\text{ot}}^{\text{ts}}$ ,

so that we have two rates

$$\begin{aligned}\lambda_{o \rightarrow t} &\propto \exp \left[ -\beta \left( E_{ot}^{\text{ts}} - E_o \right) \right] \\ \lambda_{t \rightarrow o} &\propto \exp \left[ -\beta \left( E_{ot}^{\text{ts}} - E_t \right) \right]\end{aligned}\quad (35)$$

and we have the probabilities

$$\rho_o = \frac{\lambda_{t \rightarrow o}}{2\lambda_{o \rightarrow t} + \lambda_{t \rightarrow o}}, \quad \rho_t = \frac{\lambda_{o \rightarrow t}}{2\lambda_{o \rightarrow t} + \lambda_{t \rightarrow o}}, \quad (36)$$

such that  $\rho_o + 2\rho_t = 1$ . The site vector basis for all of our sites are empty, so that we have no correlation terms to consider. From this, we have the isotropic diffusivity from Eqn. 12,

$$\begin{aligned}\underline{D} &= \left\{ 8 \frac{1}{2} \frac{a^2}{16} \lambda_{o \rightarrow t} \rho_o + 4 \cdot 2 \frac{1}{2} \frac{a^2}{16} \lambda_{t \rightarrow o} \rho_t \right\} \mathbf{1} \\ &= \frac{a^2}{2} \lambda_{o \rightarrow t} \rho_o \mathbf{1} = \frac{a^2}{2} \lambda_{t \rightarrow o} \rho_t \mathbf{1},\end{aligned}\quad (37)$$

and the isotropic activation energy from Eqn. 1 and Eqn. 22,

$$\begin{aligned}\underline{E}^{\text{act}} &= \underline{D}^{-1} \left\{ 8 \frac{1}{2} \frac{a^2}{16} \left[ E_{ot}^{\text{ts}} - \langle E \rangle \right] \lambda_{o \rightarrow t} \rho_o + 4 \cdot 2 \frac{1}{2} \frac{a^2}{16} \left[ E_{ot}^{\text{ts}} - \langle E \rangle \right] \lambda_{t \rightarrow o} \rho_t \right\} \mathbf{1} \\ &= \left[ E_{ot}^{\text{ts}} - \langle E \rangle \right] \mathbf{1}.\end{aligned}\quad (38)$$

We can also determine the elastodiffusion tensor and (isotropic) activation volume for an interstitial in FCC. To determine the elastodiffusion tensor, we note that the octahedral and tetrahedral elastic dipoles are fully isotropic by cubic and tetrahedral symmetry, respectively. The elastic dipole corresponding to the transition state has a three-fold rotation axis given by the  $\langle 111 \rangle$  vector connecting the two sites, so that the dipole has a component parallel  $P_{ot,\parallel}^{\text{ts}}$  to the rotation axis  $\mathbf{v} = \frac{1}{\sqrt{3}} \langle 111 \rangle$  and a component perpendicular  $P_{ot,\perp}^{\text{ts}}$ ,

$$\underline{P}_{ot}^{\text{ts}}(\mathbf{v}) = P_{ot,\parallel}^{\text{ts}} \mathbf{v} \otimes \mathbf{v} + P_{ot,\perp}^{\text{ts}} (\mathbf{1} - \mathbf{v} \otimes \mathbf{v}) = P_{ot,\perp}^{\text{ts}} \mathbf{1} + (P_{ot,\parallel}^{\text{ts}} - P_{ot,\perp}^{\text{ts}}) \mathbf{v} \otimes \mathbf{v} \quad (39)$$

which, in component form, is

$$\underline{P}_{ot}^{\text{ts}}([s_1, s_2, s_3])_{ab} = P_{ot,\perp}^{\text{ts}} \delta_{ab} + (P_{ot,\parallel}^{\text{ts}} - P_{ot,\perp}^{\text{ts}}) \frac{s_a s_b}{3} \quad (40)$$

where  $s_a = \pm 1$  is the sign of the corresponding component of  $\mathbf{v}$ . Similarly,  $(\delta \mathbf{x} \otimes \delta \mathbf{x})_{ab} = s_a s_b a^4 / 16$ . We can then evaluate the elastodiffusion tensor, using the Voigt notation for fourth-rank tensors, and Eqn. 34,

$$\begin{aligned}d_{11} = d_{22} = d_{33} &= a^2 \left\{ 1 + \frac{\beta}{2} \left( \frac{2}{3} P_{ot,\perp}^{\text{ts}} + \frac{1}{3} P_{ot,\parallel}^{\text{ts}} - \langle P \rangle \right) \right\} \lambda_{o \rightarrow t} \rho_o \\ d_{12} = d_{13} = d_{23} = d_{21} = d_{31} = d_{32} &= a^2 \left\{ \frac{\beta}{2} \left( \frac{2}{3} P_{ot,\perp}^{\text{ts}} + \frac{1}{3} P_{ot,\parallel}^{\text{ts}} - \langle P \rangle \right) \right\} \lambda_{o \rightarrow t} \rho_o \\ d_{44} = d_{55} = d_{66} &= a^2 \left\{ \frac{1}{2} + \frac{\beta}{2} \left( \frac{1}{3} P_{ot,\parallel}^{\text{ts}} - \frac{1}{3} P_{ot,\perp}^{\text{ts}} \right) \right\} \lambda_{o \rightarrow t} \rho_o\end{aligned}\quad (41)$$

and all other terms are zero, and where  $\langle \underline{P} \rangle = \underline{P}_o \rho_o + \underline{P}_t 2\rho_t$ . Due to symmetry, both  $\underline{P}_o$  and  $\underline{P}_t$  are isotropic (scalars). Then, the activation volume tensor from Eqn. 3 is

$$\begin{aligned} V_{11}^{\text{act}} &= 2k_B T S_{11} + \frac{1}{3} \left( \frac{1}{3} \text{Tr } \underline{P}_{\text{ot}}^{\text{ts}} - \langle \underline{P} \rangle \right) B^{-1} \\ V_{12}^{\text{act}} &= 2k_B T S_{12} + \frac{1}{3} \left( \frac{1}{3} \text{Tr } \underline{P}_{\text{ot}}^{\text{ts}} - \langle \underline{P} \rangle \right) B^{-1} \\ V_{44}^{\text{act}} &= \frac{1}{2} k_B T S_{44} + \frac{1}{6} (P_{\text{ot},\parallel}^{\text{ts}} - P_{\text{ot},\perp}^{\text{ts}}) S_{44} \end{aligned} \quad (42)$$

for bulk modulus  $B^{-1} = 3(S_{11} + 2S_{12})$ , and where  $\text{Tr } \underline{P}_{\text{ot}}^{\text{ts}} = 2P_{\text{ot},\perp}^{\text{ts}} + P_{\text{ot},\parallel}^{\text{ts}}$ . The terms proportional to  $k_B T$  all represent geometric contributions, while the remaining terms correspond to the sensitivity of the transition states and interstitial sites to stress.

### 3.2. Body-centered cubic

In a body-centered cubic crystal, the common interstitial site is an octahedral site, with the possibility of tetrahedral sites either as metastable or transition sites. Here, to compare with the face-centered cubic and hexagonal closed-packed case, we will take both octahedral and tetrahedral sites to be stable, and so there will be diffusion from octahedral to tetrahedral sites and between tetrahedral sites. The body-centered cubic space group is  $Im\bar{3}m$ , where the solvent atoms occupy  $2a$  Wyckoff positions, with octahedrals at  $6b$  and tetrahedrals at  $12d$ . [26] Unlike our other cases, there are three octahedral sites and six tetrahedral sites for every “solvent” atom; moreover, the octahedral sites do not have cubic point group symmetry, but rather tetragonal symmetry. If we consider a representative octahedral at  $[00\frac{1}{2}]$ , it connects to four tetrahedrals in the  $(001)$  plane, with jump vectors  $\pm\frac{a}{4}[100]$  and  $\pm\frac{a}{4}[010]$  and a transition state energy  $E_{\text{ot}}^{\text{ts}}$ . The elastic dipole associated with this octahedral has two components,

$$\underline{P}_o = P_{o,\perp}(\mathbf{x} \otimes \mathbf{x} + \mathbf{y} \otimes \mathbf{y}) + P_{o,p}\mathbf{z} \otimes \mathbf{z},$$

corresponding to the in-plane  $P_{o,p}$  and perpendicular  $P_{o,\perp}$  components. The transition state along  $[100]$  has an elastic dipole with three components,

$$\underline{P}_{\text{ot}}^{\text{ts}} = P_{\text{ot},\parallel}^{\text{ts}}\mathbf{x} \otimes \mathbf{x} + P_{\text{ot},\perp}^{\text{ts}}\mathbf{y} \otimes \mathbf{y} + P_{\text{ot},p}^{\text{ts}}\mathbf{z} \otimes \mathbf{z},$$

corresponding to the along-hop  $P_{\text{ot},\parallel}^{\text{ts}}$  as well as perpendicular and in-plane components. Considering next a representative tetrahedral at  $[0\frac{1}{2}\frac{1}{2}]$ , it connects to two octahedrals and four tetrahedrals, with jump vectors  $\{\frac{a}{4}[01\bar{1}], \frac{a}{4}[011], \frac{a}{4}[\bar{1}10], \frac{a}{4}[1\bar{1}0]\}$  and a transition state energy  $E_{\text{tt}}^{\text{ts}}$ . The elastic dipole associated with this tetrahedral has two components,

$$\underline{P}_t = P_{t,\perp}(\mathbf{x} \otimes \mathbf{x} + \mathbf{z} \otimes \mathbf{z}) + P_{t,a}\mathbf{y} \otimes \mathbf{y},$$

corresponding to the four-fold rotation/mirror axis  $P_{t,a}$  and perpendicular  $P_{t,\perp}$  components. The transition state along  $[01\bar{1}]$  has an elastic dipole with three components,

$$\underline{P}_{\text{tt}}^{\text{ts}} = P_{\text{tt},\parallel}^{\text{ts}} \frac{1}{2} (\mathbf{y} - \mathbf{z}) \otimes (\mathbf{y} - \mathbf{z}) + P_{\text{tt},\perp}^{\text{ts}} \frac{1}{2} (\mathbf{y} + \mathbf{z}) \otimes (\mathbf{y} + \mathbf{z}) + P_{\text{tt},p}^{\text{ts}} \mathbf{x} \otimes \mathbf{x},$$

corresponding to the along-hop as well as perpendicular and in-plane components. In order for the site probabilities to match what was found in FCC, Eqn. 36, we need to account for the increased multiplicity of octahedral and tetrahedral sites. We choose to keep the same normalization  $\rho_o + 2\rho_t = 1$ , and include a factor of 1/3 for each octahedral and tetrahedral site probability so that similarities to the face-centered cubic case can be most easily drawn.

Both the octahedral and tetrahedral sites have empty site vector bases, and so we do not need to consider correlation effects in the computation of the diffusivity or its derivatives. From this, we have the isotropic diffusivity of Eqn. 12,

$$\underline{D} = \frac{a^2}{12} \{\rho_o \lambda_{o \rightarrow t} + 2\rho_t \lambda_{t \rightarrow t}\} \mathbf{1}, \quad (43)$$

and the isotropic activation energy from Eqn. 1 and Eqn. 22,

$$\begin{aligned} \underline{E}^{\text{act}} &= \underline{D}^{-1} \left\{ \frac{a^2}{12} [E_{\text{ot}}^{\text{ts}} - \langle E \rangle] \rho_o \lambda_{o \rightarrow t} + \frac{a^2}{6} [E_{\text{tt}}^{\text{ts}} - \langle E \rangle] \rho_t \lambda_{t \rightarrow t} \right\} \mathbf{1} \\ &= \frac{[E_{\text{ot}}^{\text{ts}} - \langle E \rangle] \rho_o \lambda_{o \rightarrow t} + [E_{\text{tt}}^{\text{ts}} - \langle E \rangle] 2\rho_t \lambda_{t \rightarrow t}}{\rho_o \lambda_{o \rightarrow t} + 2\rho_t \lambda_{t \rightarrow t}} \mathbf{1}. \end{aligned} \quad (44)$$

The activation energy changes from  $E_{\text{ot}}^{\text{ts}} - \langle E \rangle$  when the octahedral to tetrahedral rate  $\rho_o \lambda_{o \rightarrow t}$  is the fastest to  $E_{\text{tt}}^{\text{ts}} - \langle E \rangle$  when the tetrahedral to tetrahedral rate  $\rho_t \lambda_{t \rightarrow t}$  is the fastest. Finally, the elastodiffusion tensor has the cubic symmetry of the lattice, and using the Voigt notation for fourth-rank tensors with Eqn. 34,

$$\begin{aligned} d_{11} = d_{22} = d_{33} &= \frac{a^2}{12} \left\{ 2(\rho_o \lambda_{o \rightarrow t} + 2\rho_t \lambda_{t \rightarrow t}) + \frac{\beta}{2} (2P_{\text{ot},\parallel}^{\text{ts}} - 2\langle \underline{P} \rangle) \rho_o \lambda_{o \rightarrow t} + \beta (P_{\text{tt},\parallel}^{\text{ts}} + P_{\text{tt},\perp}^{\text{ts}} - 2\langle \underline{P} \rangle) \rho_t \lambda_{t \rightarrow t} \right\} \\ d_{12} = d_{13} = d_{23} = d_{21} = d_{31} = d_{32} &= \frac{a^2}{12} \left\{ \frac{\beta}{2} (P_{\text{ot},\perp}^{\text{ts}} + P_{\text{ot},\text{p}}^{\text{ts}} - 2\langle \underline{P} \rangle) \rho_o \lambda_{o \rightarrow t} + \beta \left( \frac{1}{2} P_{\text{tt},\parallel}^{\text{ts}} + \frac{1}{2} P_{\text{tt},\perp}^{\text{ts}} + P_{\text{tt},\text{p}}^{\text{ts}} - 2\langle \underline{P} \rangle \right) \rho_t \lambda_{t \rightarrow t} \right\} \\ d_{44} = d_{55} = d_{66} &= \frac{a^2}{12} \left\{ (\rho_o \lambda_{o \rightarrow t} + 2\rho_t \lambda_{t \rightarrow t}) + \beta \left( \frac{1}{2} P_{\text{tt},\parallel}^{\text{ts}} - \frac{1}{2} P_{\text{tt},\perp}^{\text{ts}} \right) \rho_t \lambda_{t \rightarrow t} \right\} \end{aligned} \quad (45)$$

and all other terms are zero, and where

$$\begin{aligned} \langle \underline{P} \rangle &= \frac{1}{3} \{ \text{Tr } \underline{P}_o \rho_o + \text{Tr } \underline{P}_t 2\rho_t \} \mathbf{1} \\ &= \left\{ \left( \frac{2}{3} P_{\text{o},\perp} + \frac{1}{3} P_{\text{o},\text{p}} \right) \rho_o + \left( \frac{2}{3} P_{\text{t},\perp} + \frac{1}{3} P_{\text{t},\text{a}} \right) 2\rho_t \right\} \mathbf{1}. \end{aligned} \quad (46)$$

Due to symmetry,  $\langle \underline{P} \rangle$  is isotropic even though  $\underline{P}_o$  and  $\underline{P}_t$  are not. Then, the activation

volume tensor from Eqn. 3 is

$$\begin{aligned}
 \underline{V}_{11}^{\text{act}} &= 2k_B T S_{11} + \frac{1}{3} \left\{ P_{\text{ot},\parallel}^{\text{ts}} S_{11} + (P_{\text{ot},\perp}^{\text{ts}} + P_{\text{ot},\text{p}}^{\text{ts}}) S_{12} - \langle \underline{P} \rangle B^{-1} \right\} \frac{\rho_o \lambda_{o \rightarrow t}}{\rho_o \lambda_{o \rightarrow t} + 2\rho_t \lambda_{t \rightarrow t}} \\
 &\quad + \left\{ \frac{1}{2} (P_{\text{tt},\parallel}^{\text{ts}} + P_{\text{tt},\perp}^{\text{ts}}) (S_{11} + S_{12}) + P_{\text{tt},\text{p}}^{\text{ts}} S_{12} - \frac{1}{3} \langle \underline{P} \rangle B^{-1} \right\} \frac{2\rho_t \lambda_{t \rightarrow t}}{\rho_o \lambda_{o \rightarrow t} + 2\rho_t \lambda_{t \rightarrow t}} \\
 \underline{V}_{12}^{\text{act}} &= 2k_B T S_{12} \\
 &\quad + \left\{ (P_{\text{ot},\parallel}^{\text{ts}} + P_{\text{ot},\perp}^{\text{ts}} + P_{\text{ot},\text{p}}^{\text{ts}}) S_{11} + \left( P_{\text{ot},\parallel}^{\text{ts}} + \frac{1}{2} (P_{\text{ot},\perp}^{\text{ts}} + P_{\text{ot},\text{p}}^{\text{ts}}) \right) S_{12} - \frac{1}{3} \langle \underline{P} \rangle B^{-1} \right\} \frac{\rho_o \lambda_{o \rightarrow t}}{\rho_o \lambda_{o \rightarrow t} + 2\rho_t \lambda_{t \rightarrow t}} \\
 &\quad + \left\{ (P_{\text{tt},\parallel}^{\text{ts}} + P_{\text{tt},\perp}^{\text{ts}} + P_{\text{tt},\text{p}}^{\text{ts}}) S_{11} + \left( \frac{3}{4} P_{\text{tt},\parallel}^{\text{ts}} + \frac{3}{4} P_{\text{tt},\perp}^{\text{ts}} + \frac{1}{2} P_{\text{tt},\text{p}}^{\text{ts}} \right) S_{12} - \frac{1}{3} \langle \underline{P} \rangle B^{-1} \right\} \frac{2\rho_t \lambda_{t \rightarrow t}}{\rho_o \lambda_{o \rightarrow t} + 2\rho_t \lambda_{t \rightarrow t}} \\
 \underline{V}_{44}^{\text{act}} &= \frac{1}{2} k_B T S_{44} + \frac{1}{8} (P_{\text{tt},\parallel}^{\text{ts}} - P_{\text{tt},\perp}^{\text{ts}}) \frac{2\rho_t \lambda_{t \rightarrow t}}{\rho_o \lambda_{o \rightarrow t} + 2\rho_t \lambda_{t \rightarrow t}} S_{44}
 \end{aligned} \tag{47}$$

for bulk modulus  $B^{-1} = 3(S_{11} + 2S_{12})$ . The terms proportional to  $k_B T$  all represent geometric contributions, while the remaining terms correspond to the sensitivity of the transition states and interstitial sites to stress. There is more anisotropy in the activation volume for BCC compared with FCC, due to the lowered symmetry in the transition states.

### 3.3. Hexagonal closed-packed

In a hexagonal closed-packed crystal, there are a variety of interstitial sites available [3, 20, 27–29], but to compare with the our cubic cases, we will consider an octahedral-tetrahedral network. The hexagonal closed-packed space group is  $P6_3mmc$ , where the solvent atoms occupy  $2c$  Wyckoff positions, with octahedrals at  $2a$  and tetrahedrals at  $4f$ . [26] Similar to FCC, there is one octahedral and two tetrahedral sites for every “solvent” atom, but the connectivity is different. The octahedral sites form an interpenetrated simple hexagonal lattice, where each octahedral connects to two octahedrals directly above and below ( $\pm \frac{c}{2} [0001]$  jumps) and six tetrahedrals (either  $\{\pm \frac{a}{3} [1\bar{1}0 \frac{3c}{8a}], \pm \frac{a}{3} [01\bar{1} \frac{3c}{8a}], \pm \frac{a}{3} [\bar{1}01 \frac{3c}{8a}]\}$  or  $\{\pm \frac{a}{3} [\bar{1}10 \frac{3c}{8a}], \pm \frac{a}{3} [0\bar{1}1 \frac{3c}{8a}], \pm \frac{a}{3} [10\bar{1} \frac{3c}{8a}]\}$ ).<sup>3</sup> The tetrahedrals are each connected to one other tetrahedral (either  $\frac{c}{8} [0001]$  or  $\frac{c}{8} [000\bar{1}]$ ) and three tetrahedrals (opposing the octahedral jumps, and with the  $c$ -axis contribution negating the tetrahedral jump contribution). This introduces three transition state energies,  $E_{\text{ot}}^{\text{ts}}$ ,  $E_{\text{tt}}^{\text{ts}}$ , and  $E_{\text{oo}}^{\text{ts}}$  with corresponding rates among octahedrals and tetrahedrals. The site probabilities match what was found in FCC, Eqn. 36. From this, we can generate the anisotropic diffusivity tensor (take  $\mathbf{x}$  and  $\mathbf{y}$  to be orthonormal basis vectors in the basal plane and  $\mathbf{z}$  along the  $c$ -axis),

$$\underline{D} = \frac{c^2}{4} \mathbf{z} \otimes \mathbf{z} \rho_o \lambda_{o \rightarrow o} + 2 \left( \frac{3}{2} a^2 \mathbf{x} \otimes \mathbf{x} + \frac{3}{2} a^2 \mathbf{y} \otimes \mathbf{y} \right) \rho_o \lambda_{o \rightarrow t} + 3 \cdot 2 \left( \frac{c^2}{64} \mathbf{z} \otimes \mathbf{z} \right) \rho_o \lambda_{o \rightarrow t} + \frac{c^2}{16} \mathbf{z} \otimes \mathbf{z} \rho_t \lambda_{t \rightarrow t} + \mathbf{b}_t \otimes \boldsymbol{\gamma}_t \tag{48}$$

once we solve for the velocity and bias-correction vectors. The octahedral sites possess a threefold axis (pointing along  $\mathbf{z}$ ) and a mirror plane perpendicular to that axis, so there

<sup>3</sup>The tetrahedral sites are not required to sit at a  $c/8$  distance away from the basal plane—only that their positions be symmetric under a mirror through the basal plane. Hence, the general case would introduce the Wyckoff parameter  $z$  to character this position in the  $4f$  site. However, in the case of diffusion, only the long-range contribution matters, so that change in jump vectors due to  $z$  exactly cancels out, and we use  $z = 1/8$  for simplicity.

is no vector basis at those sites; the tetrahedral sites also possess the same threefold axis, but no mirror plane, and so a vector can point along  $\mathbf{z}$  at the tetrahedral sites. A mirror operation through the basal plane maps one tetrahedral sites to the other connected by a tetrahedral-to-tetrahedral jump; thus, the single site vector basis alternates  $\pm\mathbf{z}$  for each successive (basal) plane of tetrahedrals. In this basis, the symmetrized rate matrix only has a single entry based on the escape rate from a tetrahedral and the tetrahedral-to-tetrahedral jump,

$$\begin{aligned}\omega &= \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \cdot \begin{pmatrix} -\lambda_{t \rightarrow t} - 3\lambda_{t \rightarrow o} & \lambda_{t \rightarrow t} \\ \lambda_{t \rightarrow t} & -\lambda_{t \rightarrow t} - 3\lambda_{t \rightarrow o} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \\ &= -(2\lambda_{t \rightarrow t} + 3\lambda_{t \rightarrow o})\end{aligned}\quad (49)$$

and the velocity vector at a tetrahedral is  $\pm \left(3\lambda_{t \rightarrow o} \frac{c}{8} - \lambda_{t \rightarrow t} \frac{c}{4}\right) \rho_t^{1/2} \mathbf{z}$ . After projecting into the site vector basis, and solving for the bias-correction vector, we have

$$\mathbf{b}_t \otimes \boldsymbol{\gamma}_t = -\frac{c^2}{32} \frac{(3\lambda_{t \rightarrow o} - 2\lambda_{t \rightarrow t})^2}{3\lambda_{t \rightarrow o} + 2\lambda_{t \rightarrow t}} \rho_t \mathbf{z} \otimes \mathbf{z} \quad (50)$$

and thus the diffusivity from Eqn. 48 is

$$\underline{D} = 3a^2(\mathbf{x} \otimes \mathbf{x} + \mathbf{y} \otimes \mathbf{y})\rho_o\lambda_{o \rightarrow t} + \frac{c^2}{4} \mathbf{z} \otimes \mathbf{z} \rho_o\lambda_{o \rightarrow o} + \frac{c^2}{4} \mathbf{z} \otimes \mathbf{z} \rho_t \frac{3\lambda_{t \rightarrow o}\lambda_{t \rightarrow t}}{3\lambda_{t \rightarrow o} + 2\lambda_{t \rightarrow t}} \quad (51)$$

Computing the activation energy tensor is complicated due to the correlation terms. The uncorrelated contribution to inverse temperature derivative is

$$\begin{aligned}-\frac{dD}{d\beta} &= \frac{c^2}{4} \mathbf{z} \otimes \mathbf{z} \rho_o\lambda_{o \rightarrow o} (E_{oo}^{ts} - \langle E \rangle) + 2 \left( \frac{3}{2} a^2 \mathbf{x} \otimes \mathbf{x} + \frac{3}{2} a^2 \mathbf{y} \otimes \mathbf{y} \right) \rho_o\lambda_{o \rightarrow t} (E_{ot}^{ts} - \langle E \rangle) \\ &\quad + 3 \cdot 2 \left( \frac{c^2}{64} \mathbf{z} \otimes \mathbf{z} \right) \rho_o\lambda_{o \rightarrow t} (E_{ot}^{ts} - \langle E \rangle) + \frac{c^2}{16} \mathbf{z} \otimes \mathbf{z} \rho_t\lambda_{t \rightarrow t} (E_{tt}^{ts} - \langle E \rangle)\end{aligned}\quad (52)$$

The change in the velocity vector for the tetrahedral site is

$$\delta \mathbf{b}_t = \pm \left( \frac{3c}{8} \lambda_{t \rightarrow o} \left[ E_{ot}^{ts} - \frac{1}{2}(E_t + \langle E \rangle) \right] - \frac{c}{4} \lambda_{t \rightarrow t} \left[ E_{tt}^{ts} - \frac{1}{2}(E_t + \langle E \rangle) \right] \right) \rho_t^{1/2} \mathbf{z} \quad (53)$$

and for the symmetric rate matrix

$$\delta \omega = - \left( 2\lambda_{t \rightarrow t} [E_{tt}^{ts} - E_t] + 3\lambda_{t \rightarrow o} [E_{ot}^{ts} - E_t] \right). \quad (54)$$

After we combine all of the terms into Eqn. 22, we have an activation energy tensor

$$\begin{aligned}\underline{E}^{\text{act}} &= (E_{ot}^{ts} - \langle E \rangle) (\mathbf{x} \otimes \mathbf{x} + \mathbf{y} \otimes \mathbf{y}) + \\ &\quad \left( \rho_o\lambda_{o \rightarrow o} + \rho_t \frac{3\lambda_{t \rightarrow o}\lambda_{t \rightarrow t}}{3\lambda_{t \rightarrow o} + 2\lambda_{t \rightarrow t}} \right)^{-1} \left[ \{E_{oo}^{ts} - \langle E \rangle\} \rho_o\lambda_{o \rightarrow o} \right. \\ &\quad \left. + \left\{ \frac{3\lambda_{t \rightarrow o}}{3\lambda_{t \rightarrow o} + 2\lambda_{t \rightarrow t}} (E_{tt}^{ts} - \langle E \rangle) + \frac{2\lambda_{t \rightarrow t}}{3\lambda_{t \rightarrow o} + 2\lambda_{t \rightarrow t}} (E_{to}^{ts} - \langle E \rangle) \right\} \rho_t \frac{3\lambda_{t \rightarrow o}\lambda_{t \rightarrow t}}{3\lambda_{t \rightarrow o} + 2\lambda_{t \rightarrow t}} \right] (\mathbf{z} \otimes \mathbf{z}).\end{aligned}\quad (55)$$

For basal plane diffusivity, the activation barrier is  $(E_{\text{ot}}^{\text{ts}} - \langle E \rangle)$ , as expected. For diffusion along the  $c$ -axis, the expression in Eqn. 55 simplifies in several key limits. If  $\rho_o \lambda_{o \rightarrow o}$  is the fastest, then  $\underline{E}^{\text{act}} \approx (E_{\text{oo}}^{\text{ts}} - \langle E \rangle)$ ; diffusion is dominated by direct octahedral-to-octahedral jumps. If it is not, then the octahedral-tetrahedral network dominates; if  $\lambda_{t \rightarrow t}$  is rate-limiting (slower than  $\lambda_{t \rightarrow o}$ ),  $\underline{E}^{\text{act}} \approx (E_{\text{ot}}^{\text{ts}} - \langle E \rangle)$ , otherwise  $\underline{E}^{\text{act}} \approx (E_{\text{tt}}^{\text{ts}} - \langle E \rangle)$ . These important limits are all captured in a single expression for the activation barrier, and can be computed automatically.

The elastic dipole terms break into basal and  $c$ -axis components, with a tetragonal component for the octahedral-tetrahedral jump. The elastic dipole associated with this octahedral has two components, due to hexagonal symmetry,

$$\underline{P}_o = P_{o,b}(\mathbf{x} \otimes \mathbf{x} + \mathbf{y} \otimes \mathbf{y}) + P_{o,c} \mathbf{z} \otimes \mathbf{z},$$

corresponding to the basal  $P_{o,b}$  and  $c$ -axis  $P_{o,c}$  components. The octahedral-to-octahedral transition along the  $c$ -axis behaves similarly,

$$\underline{P}_{\text{oo}}^{\text{ts}} = P_{\text{oo},b}^{\text{ts}}(\mathbf{x} \otimes \mathbf{x} + \mathbf{y} \otimes \mathbf{y}) + P_{\text{oo},c}^{\text{ts}} \mathbf{z} \otimes \mathbf{z}.$$

The elastic dipole associated with the tetrahedral has two components,

$$\underline{P}_t = P_{t,b}(\mathbf{x} \otimes \mathbf{x} + \mathbf{y} \otimes \mathbf{y}) + P_{t,c} \mathbf{z} \otimes \mathbf{z},$$

corresponding to the basal  $P_{t,b}$  and  $c$ -axis  $P_{t,c}$  components. The tetrahedral-to-tetrahedral transition along the  $c$ -axis also has hexagonal symmetry,

$$\underline{P}_{\text{tt}}^{\text{ts}} = P_{\text{tt},b}^{\text{ts}}(\mathbf{x} \otimes \mathbf{x} + \mathbf{y} \otimes \mathbf{y}) + P_{\text{tt},c}^{\text{ts}} \mathbf{z} \otimes \mathbf{z}.$$

Finally, the octahedral-to-tetrahedral transition state does not locally have hexagonal symmetry. If we consider one jump along  $\frac{a}{3}[110\frac{2c}{8a}]$ , oriented so that it lies in the  $yz$  plane, then we can write the three components of the elastic dipole,

$$\underline{P}_{\text{ot}}^{\text{ts}} = P_{\text{ot},b}^{\text{ts}}(\mathbf{x} \otimes \mathbf{x} + \mathbf{y} \otimes \mathbf{y}) + P_{\text{ot},t}^{\text{ts}}(-\mathbf{x} \otimes \mathbf{x} + \mathbf{y} \otimes \mathbf{y}) + P_{\text{ot},c}^{\text{ts}} \mathbf{z} \otimes \mathbf{z},$$

corresponding to a basal, tetragonal, and  $c$ -axis contributions. Alternatively, we could write contributions as  $P_{\text{ot},\parallel}^{\text{ts}} = (P_{\text{ot},b}^{\text{ts}} + P_{\text{ot},t}^{\text{ts}})\mathbf{y} \otimes \mathbf{y}$  along the hop and  $P_{\text{ot},\perp}^{\text{ts}} = (P_{\text{ot},b}^{\text{ts}} - P_{\text{ot},t}^{\text{ts}})\mathbf{x} \otimes \mathbf{x}$  perpendicular to the hop. We find that, after applying the 3-fold rotation axis that the tetragonal component only appears in our bare term.

Hexagonal symmetry reduces the elastodiffusion tensor to seven independent components, of which two have contributions from correlation. The contributions from correlation come from the tetrahedral site. The change in the velocity vector for the tetrahedral site is

$$\begin{aligned} \delta \mathbf{b}_t = & \pm \left( \frac{3c}{8} \lambda_{t \rightarrow o} \left[ P_{\text{ot},b}^{\text{ts}} - \frac{1}{2}(P_{t,b} + \langle P \rangle_b) \right] - \frac{c}{4} \lambda_{t \rightarrow t} \left[ P_{\text{tt},b}^{\text{ts}} - \frac{1}{2}(P_{t,b} + \langle P \rangle_b) \right] \right) \rho_t^{1/2} (\mathbf{z} \otimes \mathbf{x} \otimes \mathbf{x} + \mathbf{z} \otimes \mathbf{y} \otimes \mathbf{y}) \\ & \pm \left( \frac{3c}{8} \lambda_{t \rightarrow o} \left[ P_{\text{ot},c}^{\text{ts}} - \frac{1}{2}(P_{t,c} + \langle P \rangle_c) \right] - \frac{c}{4} \lambda_{t \rightarrow t} \left[ P_{\text{tt},c}^{\text{ts}} - \frac{1}{2}(P_{t,c} + \langle P \rangle_c) \right] \right) \rho_t^{1/2} \mathbf{z} \otimes \mathbf{z} \otimes \mathbf{z} \quad (56) \end{aligned}$$

and for the symmetric rate matrix

$$\delta\omega = -\left(2\lambda_{t\rightarrow t}\left[P_{tt,b}^{ts} - P_{t,b}\right] + 3\lambda_{t\rightarrow o}\left[P_{tt,b}^{ts} - P_{t,b}\right]\right)(\mathbf{x} \otimes \mathbf{x} + \mathbf{y} \otimes \mathbf{y}) \\ - \left(2\lambda_{t\rightarrow t}\left[P_{tt,c}^{ts} - P_{t,c}\right] + 3\lambda_{t\rightarrow o}\left[P_{tt,c}^{ts} - P_{t,c}\right]\right)\mathbf{z} \otimes \mathbf{z}, \quad (57)$$

where  $\langle \underline{P} \rangle = \underline{P}_o \rho_o + \underline{P}_t 2\rho_t$ , as in our FCC lattice. Symmetry gives us 6 unique elastodiffusion components, as basal-plane isotropy requires  $2d_{66} = d_{11} - d_{22}$ . Moreover, only the  $d_{31} = d_{32}$  and  $d_{33}$  components have contributions from correlations; only  $d_{11} = d_{22}$ ,  $d_{33}$ ,  $d_{44} = d_{55}$  and  $d_{66}$  have contributions from the geometric terms; and  $d_{44} = d_{55}$  only has contributions from geometry. Using the Voigt notation for fourth-rank tensors with Eqn. 34,

$$d_{11} = d_{22} = a^2 \left\{ 6 + \beta \left( P_{ot,b}^{ts} + \frac{1}{2} P_{ot,t}^{ts} - \langle \underline{P} \rangle_b \right) \right\} \rho_o \lambda_{o\rightarrow t} \\ d_{33} = \frac{c^2}{4} \left\{ \left[ 2 + \beta \left( P_{oo,c}^{ts} - \langle \underline{P} \rangle_c \right) \right] \rho_o \lambda_{o\rightarrow o} \right. \\ \left. + \left[ 2 + \frac{3\lambda_{t\rightarrow o}}{3\lambda_{t\rightarrow o} + 2\lambda_{t\rightarrow t}} \left( P_{tt,c}^{ts} - \langle \underline{P} \rangle_c \right) + \frac{2\lambda_{t\rightarrow t}}{3\lambda_{t\rightarrow o} + 2\lambda_{t\rightarrow t}} \left( P_{to,c}^{ts} - \langle \underline{P} \rangle_c \right) \right] \rho_t \frac{3\lambda_{t\rightarrow o}\lambda_{t\rightarrow t}}{3\lambda_{t\rightarrow o} + 2\lambda_{t\rightarrow t}} \right\} \\ d_{12} = d_{21} = a^2 \left\{ \beta \left( P_{ot,b}^{ts} - \frac{1}{2} P_{ot,t}^{ts} - \langle \underline{P} \rangle_b \right) \right\} \rho_o \lambda_{o\rightarrow t} \\ d_{13} = d_{23} = a^2 \left\{ \beta \left( P_{ot,c}^{ts} - \langle \underline{P} \rangle_c \right) \right\} \rho_o \lambda_{o\rightarrow t} \\ d_{31} = d_{32} = \frac{c^2}{4} \left\{ \beta \left( P_{oo,b}^{ts} - \langle \underline{P} \rangle_b \right) \rho_o \lambda_{o\rightarrow o} \right. \\ \left. + \left[ \frac{3\lambda_{t\rightarrow o}}{3\lambda_{t\rightarrow o} + 2\lambda_{t\rightarrow t}} \left( P_{tt,b}^{ts} - \langle \underline{P} \rangle_b \right) + \frac{2\lambda_{t\rightarrow t}}{3\lambda_{t\rightarrow o} + 2\lambda_{t\rightarrow t}} \left( P_{to,b}^{ts} - \langle \underline{P} \rangle_b \right) \right] \rho_t \frac{3\lambda_{t\rightarrow o}\lambda_{t\rightarrow t}}{3\lambda_{t\rightarrow o} + 2\lambda_{t\rightarrow t}} \right\} \\ d_{44} = d_{55} = \frac{3a^2}{2} \rho_o \lambda_{o\rightarrow t} + \frac{c^2}{8} \rho_o \lambda_{o\rightarrow o} + \frac{c^2}{8} \rho_t \frac{3\lambda_{t\rightarrow o}\lambda_{t\rightarrow t}}{3\lambda_{t\rightarrow o} + 2\lambda_{t\rightarrow t}} \\ d_{66} = \frac{d_{11} - d_{12}}{2} = a^2 \left\{ 3 + \frac{\beta}{2} P_{ot,t}^{ts} \right\} \rho_o \lambda_{o\rightarrow t} \quad (58)$$

and all other terms are zero. The activation volume tensor can be derived as before; the scalar activation volume from Eqn. 4 is

$$V^{\text{act}} = \frac{k_B T}{9B} \left\{ \left( D_b^{-1} d_{11} + D_b^{-1} d_{12} + D_c^{-1} d_{31} \right) \frac{2S_{11} + 2S_{12} + 2S_{13}}{2S_{11} + 2S_{12} + 4S_{13} + S_{33}} \right. \\ \left. + \left( 2D_b^{-1} d_{13} + D_c^{-1} d_{33} \right) \frac{2S_{13} + S_{33}}{2S_{11} + 2S_{12} + 4S_{13} + S_{33}} \right\}, \quad (59)$$

where the bulk modulus  $B^{-1} = 2S_{11} + 2S_{12} + 4S_{13} + S_{33}$ .

### 3.4. Carbon in iron

Carbon in body-centered cubic iron sits on octahedral sites, and transitions through tetrahedral transition states. Veiga *et al.*[9] used an EAM potential to compute the octahedral and tetrahedral energies and elastic dipoles. The lattice constant for this potential



is  $a_0 = 2.8553 \text{ \AA}$ , and the elastic constants are  $C_{11} = 243 \text{ GPa}$ ,  $C_{12} = 145 \text{ GPa}$ , and  $C_{44} = 116 \text{ GPa}$ . The tetrahedral transition state is  $0.816 \text{ eV}$  above the octahedral site, and the attempt frequency is taken as  $10 \text{ THz}$ [9]. The dipole tensors can be separated into parallel and perpendicular components; the parallel direction points towards the closest Fe atoms for the C, while the perpendicular components lie in the interstitial plane. For the octahedral, the parallel component is  $8.03 \text{ eV}$ , and the perpendicular components are both  $3.40 \text{ eV}$ ; for the tetrahedral transition state, the parallel component is  $4.87 \text{ eV}$ , and the perpendicular are both  $6.66 \text{ eV}$ .

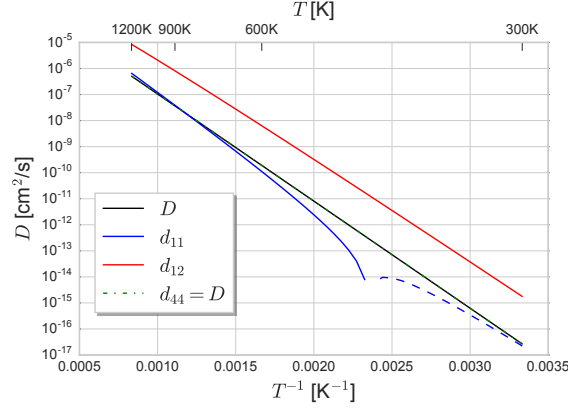


Figure 1. Diffusivity and elastodiffusivity of C in BCC Fe between 300K and 1200K. The carbon diffusivity is Arrhenius, with  $D(T) = 1.359 \times 10^{-3} \exp(-0.816/k_B T) \text{ cm}^2/\text{s}$ . The derivative  $d_{11} = dD_{xx}/d\epsilon_{xx}$  changes sign: negative below  $\approx 425\text{K}$ , and positive above. The derivative  $d_{44} = dD_{xy}/d\epsilon_{xy} = D$ , as it only has a contribution from the geometric change with strain.

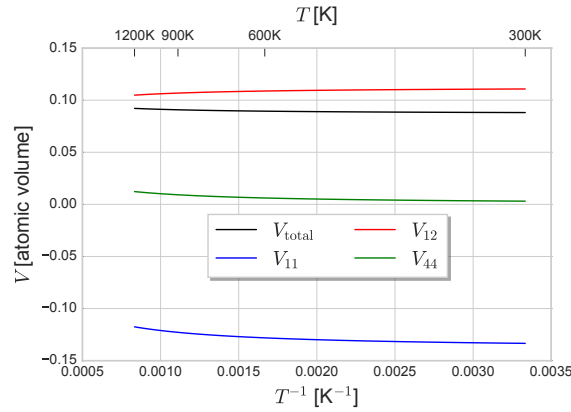


Figure 2. Activation volume of C in BCC Fe. The total activation volume from Eqn. 4 is the sensitivity to the isotropic diffusivity to hydrostatic pressure; at 300K it is  $0.0307\Omega_{\text{Fe}} = 0.357 \text{ \AA}^3$ . The components from Eqn. 3 have cubic symmetry, and so the representative terms at 300K are  $V_{11} = V_{xxx} = -0.1175\Omega_{\text{Fe}} = -1.368 \text{ \AA}^3$ ,  $V_{12} = V_{xyy} = 0.1048\Omega_{\text{Fe}} = 1.220 \text{ \AA}^3$ , and  $V_{44} = 2V_{xyxy} = 0.0122\Omega_{\text{Fe}} = 0.142 \text{ \AA}^3$ . The only contribution to  $V_{xyxy}$  is geometric.

Fig. 1 shows the Arrhenius behavior of the diffusivity, while Fig. 2 reveals unusual behavior in the activation volume from 300K to 1200K. The diffusivity follows a simple Arrhenius relationship. The elastodiffusion tensor has three unique components:  $d_{11} = dD_{xx}/d\epsilon_{xx}$ ,  $d_{12} = dD_{xx}/d\epsilon_{yy}$ ,  $d_{44} = dD_{xy}/d\epsilon_{xy}$ . The  $d_{44}$  term only contains geometric contributions, while  $d_{11}$  changes sign as the (positive) geometric contribution dominates at high temperatures and the (negative) activation contribution dominates at low temperatures. The  $d_{12}$  term is an order of magnitude larger than the other elastodiffusivity coefficients,

and also leads to the unusual activation volume behavior where  $V_{11}$  is negative,  $V_{12}$  is positive, and  $V_{\text{total}} = V_{11} + 2V_{12}$  is positive. This means that for a compressive uniaxial stress along  $\langle 100 \rangle$ , the diffusivity for carbon *increases* along  $\langle 100 \rangle$ , while the perpendicular diffusivities *decrease*. The unusual behavior is caused by the Poisson effect combined with the large  $d_{12}$  value and small  $d_{11}$  value. This still occurs while a hydrostatic compressive stress decreases the diffusivity.

#### 4. Conclusions

The general expression for the elastodiffusion tensor, activation energy and volume tensors build on previous work that assumed interstitial migration without correlation. As shown here, correlation contributions must be included even in cases of interstitial diffusion where dictated by the diffusion network. The application of this method extends beyond interstitial solutes in a crystal, but can also be used for more complex cases such as self-interstitials (which may possess multiple intermediate states) or diffusion through boundaries. There is an open-source numerical implementation of the diffusion, elastodiffusion, and activation barrier tensors, described in Appendix C[19]. The perturbation analysis laid out here can also be applied to error propagation, by calculating the derivative of diffusivity with respect to site and transition energies. It should be possible, though more complicated, to derive similar expressions for the elastodiffusion tensor for more complex diffusion mechanisms, such as vacancy-mediated solute diffusion.

#### Acknowledgements

The author thanks Maylise Nastar and Pascal Bellon for helpful conversations.

#### Disclosure

The author has no conflicts of interest.

#### Funding

Research supported in part by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering under Award #DE-FG02-05ER46217, through the Frederick Seitz Materials Research Laboratory, in part by the Office of Naval Research grant N000141210752, and in part by the National Science Foundation Award 1411106. Part of this research was performed while the author was visiting the Institute for Pure and Applied Mathematics (IPAM) at UCLA, which is supported by the National Science Foundation (NSF).

#### References

- [1] C.P. Flynn, *Point defects and diffusion*, Clarendon Press, Oxford, 1972.
- [2] A.R. Allnatt and A.B. Lidiard, *Atomic transport in solids*, chap. 5, Cambridge University Press, Cambridge (1993), p. 202.

- [3] H.H. Wu and D.R. Trinkle, *Direct diffusion through interpenetrating networks: Oxygen in titanium*, Phys. Rev. Lett. 107 (2011) (2011) 045504.
- [4] F.C. Larche and J.W. Cahn, *The effect of self-stress on diffusion in solids*, Acta metall. 30 (1982) (1982) 1835–1845.
- [5] M.J. Aziz, *Thermodynamics of diffusion under pressure and stress: Relation to point defect mechanisms*, Appl. Phys. Lett. 70 (1997) (1997) 2810.
- [6] M.S. Daw, W. Windl, N.N. Carlson, M. Laudon and M.P. Masquelier, *Effect of stress on dopant and defect diffusion in Si: A general treatment*, Phys. Rev. B 64 (2001) (2001) 045205, Available at <http://link.aps.org/doi/10.1103/PhysRevB.64.045205>.
- [7] M.J. Aziz, Y. Zhao, H.J. Gossmann, S. Mitha, S.P. Smith and D. Schiferl, *Pressure and stress effects on the diffusion of B and Sb in Si and Si-Ge alloys*, Phys. Rev. B 73 (2006) (2006) 054101.
- [8] R.G.A. Veiga, M. Perez, C.S. Becquart, C. Domain and S. Garruchet, *Effect of the stress field of an edge dislocation on carbon diffusion in  $\alpha$ -iron: Coupling molecular statics and atomistic kinetic Monte Carlo*, Phys. Rev. B 82 (2010) (2010) 054103.
- [9] R.G.A. Veiga, M. Perez, C. Becquart, E. Clouet and C. Domain, *Comparison of atomistic and elasticity approaches for carbon diffusion near line defects in  $\alpha$ -iron*, Acta mater. 59 (2011) (2011) 6963–6974.
- [10] C.H. Woo and C.B. So, *The effect of stress on point-defect diffusion in hcp metals and irradiation creep*, Philos. Mag. A 80 (2000) (2000) 1299–1318.
- [11] W.L. Chan, R.S. Averback and Y. Ashkenazy, *Anisotropic diffusion of point defects in metals under a biaxial stress field simulation and theory*, J. Appl. Phys. 104 (2008) (2008) 023502.
- [12] E.J. Savino, *Point defect-dislocation interaction in a crystal under tension*, Philos. Mag. 36 (1977) (1977) 323–340.
- [13] P.H. Dederichs and K. Schroeder, *Anisotropic diffusion in stress fields*, Phys. Rev. B 17 (1978) (1978) 2524–2536, Available at <http://link.aps.org/doi/10.1103/PhysRevB.17.2524>.
- [14] H. Jónsson, G. Mills and K.W. Jacobsen, *Nudged elastic band method for finding minimum energy paths of transitions*, in *Classical and Quantum Dynamics in Condensed Phase Simulations*, B.J. Berne, G. Ciccotti and D.F. Coker, eds., World Scientific, Singapore, 1998, p. 385.
- [15] G. Henkelman, B.P. Uberuaga and H. Jónsson, *A climbing image nudged elastic band method for finding saddle points and minimum energy paths*, J. Chem. Phys. 113 (2000) (2000) 9901–9904.
- [16] A. Janotti, M. Krčmar, C.L. Fu and R.C. Reed, *Solute diffusion in metals: Larger atoms can move faster*, Phys. Rev. Lett. 92 (2004) (2004) 085901.
- [17] M. Mantina, Y. Wang, L.Q. Chen, Z.K. Liu and C. Wolverton, *First principles impurity diffusion coefficients*, Acta mater. 57 (2009) (2009) 4102–4108.
- [18] G.H. Vineyard, *Frequency factors and isotope effects in solid state rate processes*, J. Phys. Chem. Solids 3 (1957) (1957) 121–127.
- [19] D.R. Trinkle, ONSAGER, <http://dallastrinkle.github.io/Onsager> (2016), doi://10.5281/zenodo.51413.
- [20] R. Agarwal and D.R. Trinkle, *Light element diffusion in Mg using first principles calculations: Anisotropy and elastodiffusion*, 2016. arXiv:1605.05650.
- [21] M. Glazer and G. Burns, *Space Groups for Solid State Scientists*, 3rd ed., Elsevier, 2013.
- [22] Y. Hanlumuayang, P. Gordon, T. Neeraj and D. Chrzan, *Interactions between carbon solutes and dislocations in bcc iron*, Acta mater. 58 (2010) (2010) 5481–5490, Available at <http://www.sciencedirect.com/science/article/pii/S1359645410003794>.
- [23] C. Varvenne, F. Bruneval, M.C. Marinica and E. Clouet, *Point defect modeling in materials: Coupling ab initio and elasticity approaches*, Phys. Rev. B 88 (2013) (2013) 134102, Available at <http://link.aps.org/doi/10.1103/PhysRevB.88.134102>.
- [24] T. Garnier, V.R. Manga, P. Bellon and D.R. Trinkle, *Diffusion of Si impurities in Ni under stress: A first-principles study*, Phys. Rev. B 90 (2014) (2014) 024306.
- [25] H. Kim and D.R. Trinkle, *Size misfits of point defects in aliovalently doped SrTiO<sub>3</sub> using*

- density functional theory*, Comp. Mat. Sci. 119 (2016) (2016) 41–45.
- [26] T. Hahn (ed.), *International Tables for Crystallography*, 4th ed., Vol. A, Dordrecht: Kluwer Academic, 1996.
- [27] S.C. Middleburgh and R.W. Grimes, *Defects and transport processes in beryllium*, Acta mater. 59 (2011) (2011) 7095–7103.
- [28] P. Zhang, J. Zhao and B. Wen, *Retention and diffusion of H, He, O, C impurities in Be*, J. Nucl. Mater. 423 (2012) (2012) 164–169.
- [29] A. O’Hara and A.A. Demkov, *Oxygen and nitrogen diffusion in  $\alpha$ -hafnium from first principles*, Appl. Phys. Lett. 104 (2014), 211909, Available at <http://scitation.aip.org/content/aip/journal/apl/104/21/10.1063/1.4880657>.
- [30] G.W. Stewart, *On the perturbation of pseudo-inverses, projections and linear least squares problems*, SIAM Review 19 (1977) (1977) 634–662, Available at <http://www.jstor.org/stable/2030248>.
- [31] L.W. Johnson, R.D. Riess and J.T. Arnold, *Introduction to Linear Algebra*, chap. 3, Pearson (2002), p. 199.

## Appendix A. Pseudoinverse perturbation

The expression in Eqn. 21, which is similarly used to express the first-order change in pseudoinverse of  $\omega$  with respect to strain, comes from the perturbation of the pseudoinverse and lacks terms that would correspond to the perturbative change in the null space. In this case, the simplified expression is due to the symmetry of the matrix  $\omega$ , the perturbative changes  $\delta\omega$ , and the symmetry of the vector dot-products with  $\mathbf{b}$ . If a general matrix  $A$  has pseudoinverse  $A^+$ , then we can expand the pseudoinverse  $B^+ = (A + E)^+$  up to first order in the perturbation  $E$  as

$$B^+ = A^+ - A^+(AA^+)E(A^+A)A^+ + (A^+A)^+(A^+A)E^\dagger(1 - AA^+) - (1 - A^+A)E^\dagger(AA^+)(AA^+)^+ + O(\|E\|^2) \quad (\text{A1})$$

from Eqn. (3.24) of [30]. For our case of interest,  $A = \omega$  and  $A^+ = g$  are both real symmetric matrices; moreover, we want to compute  $\mathbf{v} \cdot B^+ \cdot \mathbf{v}$  for an arbitrary vector  $\mathbf{v}$ ,

$$\begin{aligned} \mathbf{v} \cdot B^+ \cdot \mathbf{v} &= \mathbf{v} \cdot A^+ \cdot \mathbf{v} - \mathbf{v} \cdot (A^+EA^+) \cdot \mathbf{v} \\ &\quad + \mathbf{v} \cdot \left\{ (A^2)^+(A^+A)E(1 - AA^+) - \left[ (A^2)^+(A^+A)E(1 - AA^+) \right]^\text{T} \right\} \cdot \mathbf{v} + O(\|\mathbf{v} \cdot E\|^2) \\ &= \mathbf{v} \cdot A^+ \cdot \mathbf{v} - \mathbf{v} \cdot (A^+EA^+) \cdot \mathbf{v} + O(\|\mathbf{v} \cdot E\|^2) \end{aligned} \quad (\text{A2})$$

which, for our problem, reduces to first order in  $\delta\omega$

$$\mathbf{b} \cdot \delta g \cdot \mathbf{b} = -\mathbf{b} \cdot g \delta\omega g \cdot \mathbf{b} = -\gamma \cdot \delta\omega \cdot \gamma. \quad (\text{A3})$$

## Appendix B. Vector and tensor bases at sites

To construct the vector and tensor bases, we first must determine the eigenvectors of each group operation. There are only ten types of operations available in a crystal: identity, a 2-, 3-, 4-, or 6-fold axis, mirror through a plane, or a 2-, 3-, 4-, or 6-fold axis combined with a mirror through that same plane. These can be identified from the determinant—which is

−1 if the operation contains a mirror operation or 1 if it does not—and the trace,

det	Tr	operation	det	Tr	operation
1	3	identity	−1	1	mirror
1	2	6-fold axis	−1	0	6-fold axis + mirror
1	1	4-fold axis	−1	−1	4-fold axis + mirror
1	0	3-fold axis	−1	−2	3-fold axis + mirror
1	−1	2-fold axis	−1	−3	inversion

(B1)

The eigenvalues and eigenvectors for these cases are all straightforward. The trivial cases are identity—triply-degenerate eigenvalues of 1, for the three eigenvectors  $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ —inversion—triply-degenerate eigenvalues of −1, for the three eigenvectors  $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ —and mirror—one eigenvalue of −1 with eigenvector  $\mathbf{v}$  corresponding to the mirror plane normal and double-generate eigenvalues of 1 for the two orthonormal vectors  $\mathbf{v}'$  and  $\mathbf{v}''$  that are orthogonal to  $\mathbf{v}$ . For the remaining cases, the rotational axis  $\mathbf{v}$  is always an eigenvector with an eigenvalue +1 if there is not a mirror operation, or −1 if there is a mirror operation. The remaining eigenvalues are two  $n^{\text{th}}$  roots of unity  $e^{\pm i2\pi/n}$  for the  $n$ -fold rotation axis. If  $\mathbf{v}'$  and  $\mathbf{v}''$  are two orthonormal vectors that are also orthogonal to the rotational axis  $\mathbf{v}$ , then the eigenvectors are  $(\mathbf{v}' \pm i\mathbf{v}'')/\sqrt{2}$  for the positive and negative roots of unity.

We construct vector and tensor bases that are compatible with a set of point group operations from the eigenvectors of the group operations. The bases are such that they remained unchanged under application of every group operation. For each group operation,  $d$ -rank tensors can be constructed from different dyad products of eigenvectors of the group operation. In particular, the *product* of the eigenvalues of the eigenvectors needs to be 1. Hence, for a vector basis (a 1-rank tensor), we have: (a) identity produces a 3-dimensional basis of  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ , (b) mirror produces a 2-dimensional basis of  $\{\mathbf{v}', \mathbf{v}''\}$ , two orthonormal vectors in the mirror plane, (c)  $n$ -fold rotational axes without mirror produces a 1-dimensional basis of  $\{\mathbf{v}\}$ , the rotational axis vector, and (d) all other cases produce a 0-dimensional null basis,  $\emptyset$ . Each group operation produces its own basis, and these vector bases are then intersected to produce the final basis, using some simple rules: if  $p$ - and  $q$ -dimensional bases are intersected, the resulting dimension of the final basis is not greater than  $\min\{p, q\}$ . Then, the remaining non-trivial cases are: two 2-dimensional, a 2- and 1-dimensional, and two 1-dimensional intersections. In the first case, if  $\mathbf{v}$  and  $\mathbf{v}'$  are the vectors made by the cross-product of the two vectors in each basis, then the resulting intersection will be either (a) the same 2-dimensional basis if  $\mathbf{v}$  is parallel to  $\mathbf{v}'$ , or (b) a 1-dimensional basis consisting of  $\mathbf{v} \times \mathbf{v}'$ , otherwise. In the second case, if  $\mathbf{v}_1$  is the basis for the 1-dimensional basis, and  $\mathbf{v}_2$  is the cross-product of the two vectors in the 2-dimensional basis, then the resulting intersection will be either (a) 1-dimensional basis consisting of  $\mathbf{v}_1$  if  $\mathbf{v}_1$  is perpendicular to  $\mathbf{v}_2$ , or (b) the 0-dimensional empty basis  $\emptyset$ , otherwise. Finally, if  $\mathbf{v}$  and  $\mathbf{v}'$  are the two 1-dimensional bases, the resulting intersection will either be (a)  $\mathbf{v}$  if  $\mathbf{v}$  is parallel to  $\mathbf{v}'$ , or (b) the 0-dimensional empty basis  $\emptyset$ , otherwise.

The construction of the second-rank tensor bases is more complicated, but reduces to three distinct cases. We are specifically interested in *symmetric* second-rank tensors, so the largest the bases can be is 6-dimensional. For identity and inversion, the full 6-dimensional basis is invariant:  $\{\mathbf{e}_1 \otimes \mathbf{e}_1, \mathbf{e}_2 \otimes \mathbf{e}_2, \mathbf{e}_3 \otimes \mathbf{e}_3, (\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1)/\sqrt{2}, (\mathbf{e}_1 \otimes \mathbf{e}_3 + \mathbf{e}_3 \otimes \mathbf{e}_1)/\sqrt{2}, (\mathbf{e}_2 \otimes \mathbf{e}_3 + \mathbf{e}_3 \otimes \mathbf{e}_2)/\sqrt{2}\}$ . For a mirror operation or a 2-fold rotation, if  $\mathbf{v}$  is the mirror plane normal / rotational axis, and  $\mathbf{v}'$  and  $\mathbf{v}''$  are the remaining orthonormal vectors, the 4-dimensional basis is invariant:  $\{\mathbf{v} \otimes \mathbf{v}, \mathbf{v}' \otimes \mathbf{v}', \mathbf{v}'' \otimes \mathbf{v}'', (\mathbf{v}' \otimes \mathbf{v}'' + \mathbf{v}'' \otimes \mathbf{v}')/\sqrt{2}\}$ . Finally, for the  $n$ -fold rotations (with or without a mirror), if  $\mathbf{v}$  is the rotational axis, and  $\mathbf{v}'$  and  $\mathbf{v}''$  are the remaining

orthonormal vectors, the 2-dimensional basis is invariant:  $\{\mathbf{v} \otimes \mathbf{v}, (\mathbf{v}' \otimes \mathbf{v}' + \mathbf{v}'' \otimes \mathbf{v}'')/\sqrt{2}\}$ . To intersect any two bases, the simplest approach is to find the null-space of the combined column spaces constructed from the bases using singular-value decomposition[31].

### Appendix C. Implementation

A full numerical implementation of the algorithms in Python described are available on github[19] under the MIT License. This includes algorithms to analyze a given crystal (lattice and atomic basis), find generators for the space group operations, determine all point group operations for each site, identify Wyckoff positions, generate vector and tensor bases, construct a jump network in the crystal, and identify unique jumps. Once the energies, prefactors, and elastic dipoles are determined for the unique sites and jumps, the numerical implementation can compute the diffusivity, elastodiffusion, and activation barrier tensors for a given temperature. The implementation includes numeric examples of the analytic results from this paper, with input files for FCC, BCC, and HCP interstitial diffusion.