Scalable kernel polynomial method for calculating transition rates

Chen Huang, Arthur F. Voter, and Danny Perez

Phys. Rev. B **87**, 214106 — Published 24 June 2013

DOI: 10.1103/PhysRevB.87.214106

A Scalable Kernel Polynomial Method for Calculating Transition Rates

Chen Huang, Arthur F. Voter,* and Danny Perez†
Theoretical Division T-1, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

We present an efficient method for calculating the prefactors of harmonic transition state theory rates. We reformulate the prefactors in terms of the density of states (DOS) of the Hessian matrices at the basin minimum and the saddle point. The DOS is then approximated using the kernel polynomial method as an expansion in terms of Chebyshev polynomials. The cost of the calculation scales linearly with the number of atoms, in contrast with the cubic scaling of the direct method. This approach hence greatly facilitates the investigations of kinetic processes in very large systems. We demonstrate the method by calculating the prefactors of the transition rates for two processes in bulk silver: vacancy hopping and Frenkel pair formation.

I. INTRODUCTION

Transition rates of kinetic processes in materials are essential ingredients in our understanding and prediction of material properties. For example, the dislocation nucleation rate controls yield, and hence determines how much strain pristine materials can sustain. Similarly, thermal creep is limited by dislocation climb rates, and pinning/depinning of dislocations controls plastic deformation in materials that contain defects,[1] for example in the Portevin-LeChatelier effect[2]. Computational capabilities enabling calculations of rates for various processes occuring in materials is therefore of prime importance to model materials behavior on long timescales.

While rates can be systematically computed at different levels of approximation for small systems, unfavorable scaling limits the system sizes that can be investigated. For example, a popular formalism to estimate transition rates is the harmonic approximation to transition state theory (HTST).[3] In HTST, the potential energy surfaces near the initial minimum and the saddle point for the transition are approximated by harmonic potentials, which allows for explicit calculation of the rate. However, this calculation involves the diagonalization of the Hessian matrices involved in the harmonic expansions, which entails a computational cost that scales cubically with the number of atoms in the system. In cases where large simulation cells are required, e.g., for transitions that strongly couple to long-range strain fields, such as dislocation nucleation, glide, or climb[4–7], massively parallel calculations,[7] or sampling-based approaches, such as thermodynamic integration,[8] have to be employed. Alternatively, other levels of theory can be used to estimate the free energy barrier for the transition e.g., using the finite-temperature string method[9] or umbrella sampling[10], in order to get at the transition rate.

In this work, we present a scalable approach for calculating HTST rates using the kernel polynomial method (KPM), a powerful technique developed to compute properties of large matrices in a range of settings.[11–15] In order to avoid diagonalizing Hessian matrices, we reformulate the transition rate in terms of density of states (DOS) which are then expanded in terms of Chebyshev polynomials using the KPM. The coefficients in the expansion, i.e., the moments, are obtained by stochastic sampling of the trace of appropriate polynomials. The efficient implementation of this technique enables the calculation of HTST rates at a cost that scales only linearly with the number of atoms in the system.

The paper is organized as follows. In Sec. II, we formulate HTST within the KPM formalism. We discuss the scaling of the algorithm and propose different strategies to speed up convergence. The accuracy and efficiency of the method are demonstrated in Sec. IV by the computation of the prefactors for vacancy hopping and Frenkel pair formation in bulk silver.

II. THEORY

One of the most common formalisms used to estimate transition rates in the canonical ensemble is transition state theory (TST)[16–19]. In TST, the transition rate between two states is given by the canonical expectation of the flux through the dividing surface (defined in configuration space) between these two states. The main assumption that underpins TST is that each crossing of the dividing surface corresponds to a reactive event, i.e., that every trajectory that crosses the dividing surface from the initial state will commit to the final state. In general, this does not have to be the case as so-called correlated recrossing, where a trajectory recrosses the dividing surface before having committed to the final state, does occur in reality; TST rates are therefore upper bounds on true transition rates. The missing factor, the dynamical correction factor[20,21], can be computed by simulating dynamical trajectories launched from the dividing surface. In the solid state, that correction factor is often close to unity, so that TST alone can provide a very good approximation to the true transition rate.

Within TST, the transition rate is given by:

$$k^{\mathrm{TST}} = \frac{k_B T}{h} \frac{Z_{\ddagger}}{Z_{\mathrm{min}}} \tag{1}$$

where Z_{min} is the partition function of the (initial) basin and $Z_{\ddagger}$ is the partition function of the dividing surface that separates the initial and final basins. While formally

simple, this definition is seldom used directly because of the high cost of numerically estimating partition functions with high accuracy.

A simple approximation to the partition function can be obtained by expanding the Hamiltonian of the system to second order around a local mechanical equilibrium state. The system then reduces to an ensemble of uncoupled harmonic oscillators and the partition functions can be evaluated explicitly as:

$$Z_{\min} = \frac{\exp[-\beta V_{\min}]}{\prod_{i=1}^{D} \beta\hbar\omega_{\min,i}} \tag{2}$$

$$Z_{\ddagger} = \frac{\exp[-\beta V_{\ddagger}]}{\prod_{i=1}^{D-1} \beta\hbar\omega_{\ddagger,i}} \tag{3}$$

where $V_{\min}$ and $V_{\ddagger}$ are the potential energies at the minimum and the saddle point for the transition, respectively, and the ω are the real, positive, vibrational angular frequencies of the phonon modes of the system. Here D is the total number of vibrational degrees of freedom (excluding free translations and rotations). The angular frequencies are the square-roots of the eigenvalues λ of the mass-weighted Hessian matrix $H_{i,j} = (\partial^2 V/\partial x_i \partial x_j)/\sqrt{m_i m_j}$ that locally describes the curvature of the potential.

Inserting into Eq. 1, one obtains

$$k^{\mathrm{HTST}} = \frac{1}{2\pi} \frac{\prod_{i=1}^{D} \omega_{\min,i}}{\prod_{i=1}^{D-1} \omega_{\ddagger,i}} e^{-\beta\Delta V}, \tag{4}$$

where $\Delta V = V_{\ddagger} - V_{\min}$ is the energy barrier for the transition, or in terms of the eigenvalues of the mass-weighted Hessian matrices:

$$k^{\mathrm{HTST}} = \left[\frac{1}{2\pi} \sqrt{\frac{\prod_{i=1}^{D} \lambda_{\min,i}}{\prod_{i=1}^{D-1} \lambda_{\ddagger,i}}} \right] e^{-\beta\Delta V}. \tag{5}$$

This is the celebrated harmonic approximation to TST.[3] In the last expression, the term in square brackets is the so-called prefactor. The calculation of the energy barrier for a given transition can be carried out with a number of well established and scalable methods, such as the dimer method[22] or the nudged elastic band method[23]. In the following, we therefore focus exclusively on the calculation of the prefactor.

While formally simple, the evaluation of Eq. 4 requires the diagonalization of $D \times D$ matrices, which entails a computational cost proportional to D^3. This unfavorable scaling limits the size of systems that can be investigated in practice. To overcome this limitation, we reformulate the problem in terms of a quantity that can be approximated in a scalable way. Consider the product of the real positive eigenvalues of a matrix. Taking the natural logarithm, one gets

$$\begin{aligned} \log \prod_j \lambda_j = \sum_j \log \lambda_j &= \int_{0^+}^{+\infty} d\lambda \sum_j \delta(\lambda - \lambda_j) \log \lambda \\ &= \int_{0^+}^{+\infty} d\lambda \rho(\lambda) \log \lambda. \end{aligned} \tag{6}$$

This expression now involves an integral over the eigenvalue density function $\rho(\lambda)$, i.e. the DOS, of the Hessian matrix. The key is to accurately approximate the DOS in such a way that diagonalization of H is not required.

A possible approach is to expand the DOS in terms of a rapidly convergent series, for example using orthogonal polynomials; on bounded domains, Chebyshev polynomials $T_n(x)$ are a common choice. For matrix arguments, this corresponds to the so-called KPM[11]. Without loss of generality, we first rescale H so that its spectrum falls within (-1,1). This can be done simply by forming $\tilde{H} = (H - bI)/a$ with $a = (\lambda_{\max} - \lambda_{\min})/(2 - \epsilon)$, $b = (\lambda_{\max} + \lambda_{\min})/2$, where ϵ is a small positive number. The DOS of this rescaled matrix can then be expanded in Chebyshev polynomials as:

$$\tilde{\rho}(x) = \frac{1}{\pi\sqrt{1-x^2}} \left[\mu_0 + 2\sum_{n=1}^{\infty} \mu_n T_n(x) \right], \tag{7}$$

where the moments μ_n are given by the trace of the matrix-valued Chebyshev polynomials $\mathrm{Tr}[T_n(\tilde{H})]$.[24–27]

A direct evaluation of the traces would be expensive, but they can be approximated efficiently through a stochastic evaluation of the form[24,26,28]:

$$\mathrm{Tr}[T_n(\tilde{H})] \simeq \frac{1}{R} \sum_{r=1}^{R} \langle r | \, T_n(\tilde{H}) \, | r \rangle \, . \tag{8}$$

Here $\{|r\rangle\}$ is a set of random vectors that satisfy $\langle\langle \xi_{rj} \rangle\rangle = 0$ and $\langle\langle \xi_{ri} \xi_{r'j} \rangle\rangle = \delta_{rr'}\delta_{ij}$, where $\xi_{ri} \in \mathbf{R}$ denotes the i-th element of a vector $|r\rangle$ and $\langle\langle ... \rangle\rangle$ denotes statistical average with respect to different realizations of random vectors. To be more specific,the average $\langle\langle \xi_{ri}\xi_{r'j} \rangle\rangle$ corresponds to $\lim_{R\to\infty} \frac{1}{R^2} \sum_{r,r'=0}^{R-1} \xi_{ri}\xi_{r'j}$ where the sum extends over different realizations of random vectors. In the present context, Dirac's bra/ket notation denotes the usual matrix/vector and vector/vector products

An efficient evaluation of Eq. 8 relies on two key properties. First, one does not need to form $T_n(\tilde{H})$ explicitly, but simply has to compute its projection on a vector. Second, Chebyshev polynomials possess a three-term recurrence relation, which for matrix arguments takes the form:

$$T_n(\tilde{H})|r\rangle = 2\tilde{H}T_{n-1}(\tilde{H})|r\rangle - T_{n-2}(\tilde{H})|r\rangle. \tag{9}$$

Computing the product of T_n on a vector thus reduces to a sequence of products with $\tilde{H}$. In turn, this can be efficiently computed with finite differences as (for

monoatomic systems)

$$\tilde{H}|y\rangle = \frac{(H - bI)|y\rangle}{a} = \frac{\frac{1}{m\xi}(\vec{g}(\vec{x} + \xi\vec{y}) - \vec{g}(\vec{x})) - b|y\rangle}{a}, \tag{10}$$

where $\vec{g}$ is the gradient of the potential energy evaluated at $\vec{x}$, which is the coordinate of either the minimum or the saddle point. The whole process can therefore be carried out using only interatomic forces, each of which can be obtained at a cost that scales as $O(D)$ for short-ranged potentials. ξ is a small parameter ($10^{-7}\mathring{A}$ in the present study). In this work, we varied ξ from 10^{-2} to 10^{-8} to study the dependence of prefactors on ξ. We found that prefactors converged very quickly once ξ was smaller than 10^{-4}. We also tested the centered difference formula for evaluating Eq. 10, and found that the improvement in accuracy is marginal for small enough ξ. The interatomic potential used in the present study is evaluated using interpolation arrays; the dependence on ξ may be different (probably less sensitive) for an analytical form of potential.

When computing traces with Eq. 8, one should only consider contributions from eigenmodes that have positive eigenvalues. Therefore, we need to remove these unwanted modes, such as the unstable mode at the saddle point or the free translation (rotation) modes that arise when the energy is translation (rotation) invariant. Before performing the trace in Eq. 8, the vectors $|r\rangle$ are orthogonalized against the subspace spanned by these unwanted modes, i.e.,

$$|r\rangle = |r\rangle - \sum_i \langle r|u_i\rangle |u_i\rangle,$$

where the $\{|u_i\rangle\}$ form an orthogonal basis of that subspace, and can be obtained using a Gram-Schmidt procedure. Note that orthogonality against this subspace might be lost during the calculation of the different moments due to numerical errors; in practice such orthogonalization is therefore repeated after each application of $\tilde{H}$ on a vector.

Coming back to the original problem, Eq. 6 becomes:

$$\log \prod_j \lambda_j = \int_{-1}^{1} dx \tilde{\rho}(x) \log(ax + b). \tag{11}$$

We then have:

$$\begin{aligned} &\log\left[\frac{\prod_{j=1}^{D} \lambda_{\min,j}}{\prod_{j=1}^{D-1} \lambda_{\ddagger,j}}\right] \\ &= \int_{-1}^{1} dx \frac{\log(ax+b)}{\pi\sqrt{1-x^2}} \left(\Delta\mu_0 + 2\sum_{n=1}^{\infty} \Delta\mu_n T_n(x)\right) \\ &= c_0\Delta\mu_0 + 2\sum_{n=1}^{\infty} c_n \Delta\mu_n, \end{aligned} \tag{12}$$

where we have assumed that same scaling (i.e., the same a and b) for the Hessian matrices at the minimum and the saddle point so that both spectra of the scaled matrices are contained within $[-1, 1]$. c_n and $\Delta\mu_n$ in Eq. 12 are given by:

$$c_n = \int_{-1}^{1} dx \frac{\log(ax+b)}{\pi\sqrt{1-x^2}} T_n(x), \tag{13}$$

$$\begin{aligned} \Delta\mu_n &= \mu_{\min,n} - \mu_{\ddagger,n} \\ &= \mathrm{Tr}\left[T_n(\tilde{H}_{\min}) - T_n(\tilde{H}_{\ddagger})\right], \end{aligned} \tag{14}$$

from which the HTST prefactor is directly obtained.

To evaluate Eq. 14, one could sample the traces of $T_n(\tilde{H}_{\min})$ and $T_n(\tilde{H}_{\ddagger})$ separately. Due to the stochastic nature of the evaluation, the variance of the estimates of $\Delta\mu$ will in general be larger than that of each of the terms. As demonstrated below, using matched random numbers (the same vectors $|r\rangle$) to estimate both $\mu_{\min,\mathrm{n}}$ and $\mu_{\ddagger,n}$ leads to a significant reduction of the variance in estimating $\Delta\mu_n$,[29] and hence of the computational cost. This is caused by the positive covariance between the samplings of $\mu_{\min,\mathrm{n}}$ and $\mu_{\ddagger,n}$.

The random vectors can be further constrained to exactly impose certain conditions. For example, $\Delta\mu_0$ should be exactly equal to 1, as it corresponds to the norm of the difference in DOS between the minimum and saddle points. This condition can be obeyed exactly by normalizing the random vectors to D and $D-1$ at the minimum and saddle point, respectively. This procedure is known to also significantly reduce sampling errors.[30]

In practice, the sum in Eq. 12 is truncated at a finite number of moments (M), usually on the order of hundreds, depending on the complexity of the DOS difference between the minimum and the saddle point. In the regime where the prefactor depends weakly on system size, changes in DOS, and hence in moments, can also be expected to be small. Therefore, the required number of moments, is not expected to signficantly increase with system size, for large enough systems. The same is true for the number of random numbers (R, typically of the order of a thousand) required to converge the results to a desired accuracy. Therefore, the cost of our method is expected to scale as $O(DMR)$, i.e., linearly with respect to the number of atoms in the system, which is a significant improvement over the $O(D^3)$ scaling of the direct method. Below, we give numerical evidence that this scaling is indeed achieved in practice and hence that the method can be used to efficiently investigate very large systems.

III. NUMERICAL DETAILS

In the following, we demonstrate the method for two processes in bulk silver: vacancy hopping and Frenkel pair formation. While we don't expect the prefactors for these transitions to be strongly size dependent, they

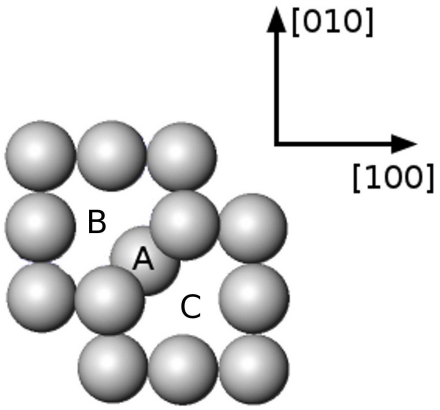

FIG. 1. Saddle point configuration for vacancy hopping in bulk silver. Atom A is moving to vacancy B and creating a new vacancy C. For clarity, only non-FCC atoms are shown (as determined by a common-neighbor analysis).

are simple and well understood processes, and hence provide good test cases for our method.The interactions between atoms are modeled with an EAM potential taken from Ref. 31. Saddle points are located using the string method.[32] The unstable mode at the saddle point and the largest and smallest eigenvalues of the Hessian matrices (required for rescaling the Hessians) are obtained with the Lanczos method[33], whose cost also scales linearly with respect to the number of atoms. The elements of the random vectors $|r\rangle$, i.e., normal random numbers with zero mean and unit variance, are generated with the Box-Muller scheme[34], in which uniform random numbers are generated by the RAN3 algorithm[35]. The coefficients c_n (Eq. 13) are integrated with the quadpack code[36]. The Jackson kernel[11,37,38] is used to regularize the expansion.

IV. RESULTS

A. Vacancy hopping

We first demonstrate our method by computing the rate prefactor for vacancy hopping in bulk silver. The saddle point configuration is shown in Figure 1: atom A is moving towards vacancy B, leaving a new vacancy C behind. The transition can therefore be seen as an exchange of the position of the vacancy with that of one of its neighbors. The energy barrier for this transition is 0.65 eV.

We first explore the impact of the properties of the random vectors on the quality of the results. We compare four distinct schemes: (1) exact zero-th moments and matched random vectors, (2) exact zero-th moments and non-matched random vectors, (3) inexact zero-th moments and matched random vectors, and (4) inexact zeroth-th moments and non-matched random vectors.

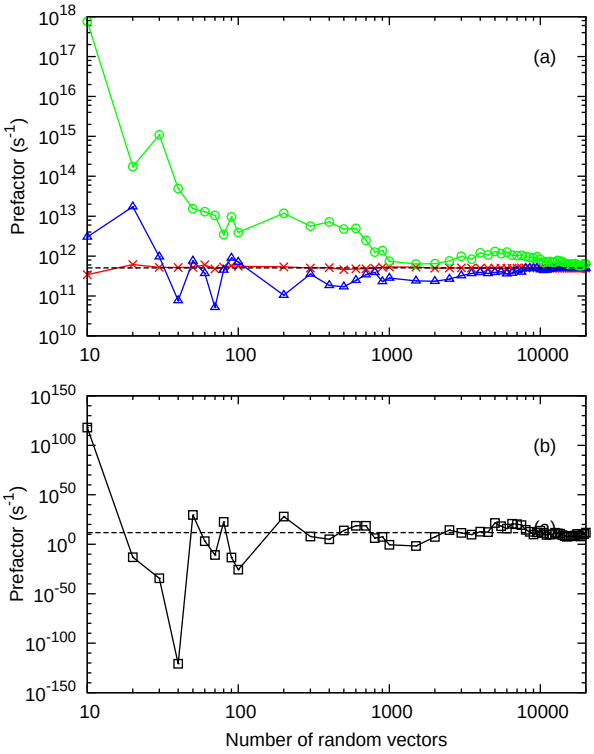

FIG. 2. Comparison of the convergence of prefactors for the vacancy hopping with respect to the number of random vectors for a system of $5 \times 5 \times 5$ unit cells, for four different schemes: (a) exact zero-th moments and matched random vectors (red crosses); exact zero-th moments and non-matched random vectors (blue triangles); inexact zero-th moments and matched random vectors (green circles). (b) Inexact zero-th moments and non-matched random vectors. Exact prefactors are shown by black dashed lines. All results are calculated with 400 moments.

The results, shown in Fig. 2, indicate that unnormalized and unmatched random vectors yield very poor results: the prefactor varies wildly, by many orders of magnitude, as the number of random vectors is increased. In contrast, imposing normalization and matching random vectors significantly improve the convergence rate. Taken together, they enable a very fast convergence: the correct order of magnitude is obtained with fewer than ten random vectors and the fluctuation with increasing the number of random vectors is modest. Unless noted otherwise, all results in this work are obtained with random vectors that are normalized and matched.

The convergence with respect to the number of moments is extremely rapid for the vacancy hopping. As shown in Figure 3, the prefactors converge to within 10% error (with respect to their converged values) with fewer than ten moments, and within 1% error with fewer than 40 moments, for all cell sizes. A similar trend is also observed for the convergence with respect to the number of random vectors. As shown in Fig. 4, about 1000 vectors are required to obtain an error at the 10% level, indepen-

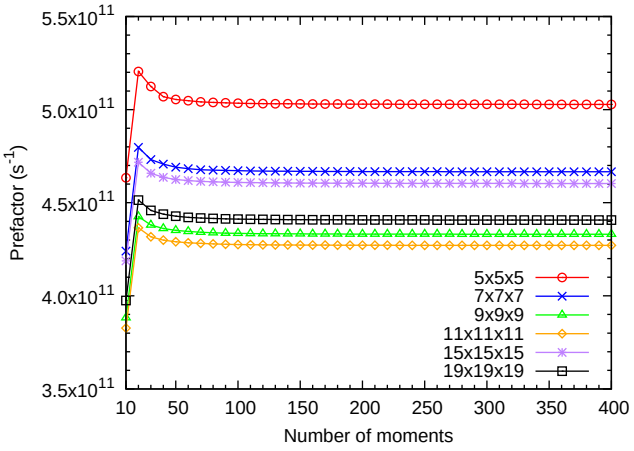

FIG. 3. Convergence of the prefactors for the vacancy hopping (using 20000 random vectors) with respect to the number of moments, for five different cell sizes. With 10, 20, 30, 40, and 50 moments, the prefactors converge to 8%, 4%, 2%, 0.8%, and 0.6% error, respectively.

dently of system size. As neither the required number of random vectors, nor the number of moments scales with system size, the computational cost of the method is linear in number of atoms, in agreement with the heuristic arguments discussed above.

Figure 5 compares the prefactors calculated from our method (with 400 moments and 20000 random vectors) with the results obtained from direct diagonalization of the Hessian matrices. The agreement between our method and the benchmark is excellent, with the errors below 1.2%. Due to high computational costs, prefactors for the $19 \times 19 \times 19$ cell were not obtained with direct diagonalization. It is interesting to note that even for such a simple transition, very large cells are required to converge the results to within 1%, as the prefactor varies by about 10% from the $5 \times 5 \times 5$ cell to the $11 \times 11 \times 11$ cell.

The convergence of the prefactor with respect to the number of moments can be understood from the behavior of the $|c_n|$ and $\Delta\mu_n$, as shown in Figure 6 for the $15 \times 15 \times 15$ cell. The results are obtained using 20000 random vectors. Perhaps contrary to intuition, convergence of the prefactors do not follow from the fast decay of the expansion coefficients (the $\Delta\mu_n$). Instead, these coefficients fluctuate around 0 with no visible sign of convergence. This can be understood from the fact that the DOS are not smooth functions but sums of delta functions. Expanding $\delta(x)$ in Chebyshev polynomials yields $\mu_n = \cos(n\pi/2)$, which is either -1, 1, or 0, depending on n. Therefore, the moments of the DOS themselves are not expected to decay with increasing order. On the other hand, the coefficients c_n correspond to the integral of a Chebyshev polynomial (a strongly oscillatory function at large n) with a smooth function. Consequently, the magnitude of c_n does decay rapidly with increasing order, as shown in Figure 6. The rapid convergence of the prefactor therefore stems from the rapid decay of the c_n combined with the alternating sign of the $\Delta\mu_n$.

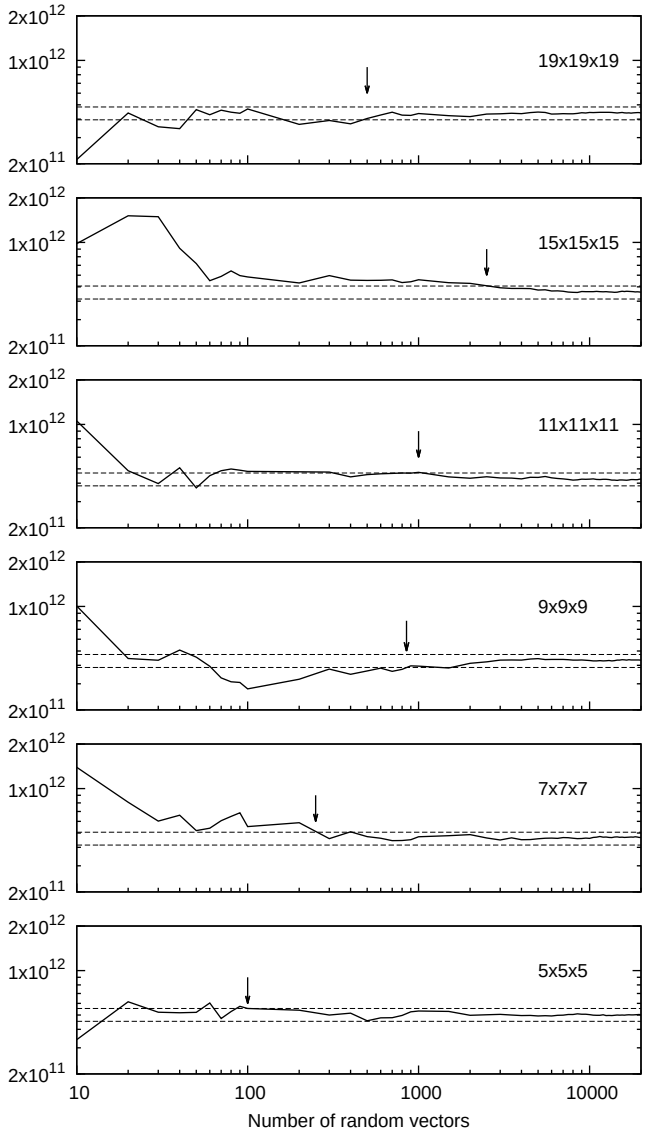

FIG. 4. Convergence of the prefactors (s^{-1}) for the vacancy hopping (using 400 moments) with respect to the number of random vectors, for five different cell sizes. The dashed horizontal lines show +/- 10% error windows from the benchmarks obtained by directly diagonalizing the Hessian matrices (except for the $19\times19\times19$ cell, where the benchmark is taken from the converged KPM results at 20000 random vectors). Arrows indicate convergence to within 10% error.

B. Frenkel pair formation

We now investigate a more complex transition, namely the nucleation of a Frenkel (vacancy-interstitial) pair from a perfect bulk crystal. The saddle point configuration and the fully formed Frenkel pair are shown in Figure 7. The pair contains a vacancy (marked "V") and

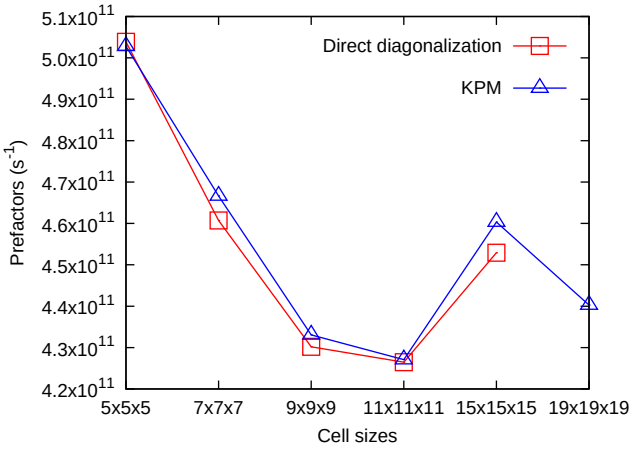

FIG. 5. Comparison of the prefactors for vacancy hopping between the KPM (with 400 moments and 20000 random vectors) and the direct diagonalization for different cell sizes. The relative errors between the KPM and the direct diagonalization are 0.2%, 1.2%, 0.7%, 0.1%, and 0.3% for cell sizes from $5 \times 5 \times 5$ to $15 \times 15 \times 15$.

a dumbbell interstitial defect (marked "D"). In this case, the energy barrier for the transition is 4.4 eV.

The convergence of the prefactor with respect to moments is shown in Figure 8 for five different cell sizes and 40000 random vectors. For all cell sizes, the prefactor estimate peaks at ~ 20 moments, and then monotonously converges. A relative error of 1% here requires around 220 moments, significantly more that the 40 that were required in the case of vacancy hopping. However, as in the former case, this number is independent of cell size.

The convergence of the prefactor (using 400 moments) with respect to the number of random vectors is shown in Figure 9, for five different cell sizes. As marked by the arrows, the number of random vectors required for convergence to within 10% error again does not increase with cells size, and the method's cost scales linearly with the number of atoms.

In Figure 10, we demonstrate the accuracy of our method compared to the direct diagonalization of Hessian matrices. The prefactors are obtained with 400 moments and 40000 random vectors. For all cell sizes, the relative errors are less than 5%, again indicating that the method provides very accurate estimates of the prefactors.

The fact that the prefactor of the Frenkel pair formation rate requires more moments than that of the vacancy hop can be traced back to characteristic of their DOSs. As seen by comparing Figures 11 and 12, which show their DOSs at the minima and saddle points, the difference in DOS is generally smoother for the vacancy hop than for the Frenkel pair formation. Further, a large number of new peaks appear in the high energy section of the saddle point's DOS for the Frenkel pair formation, as compared to only four new peaks for the vacancy hop. This is consistent with the observation that

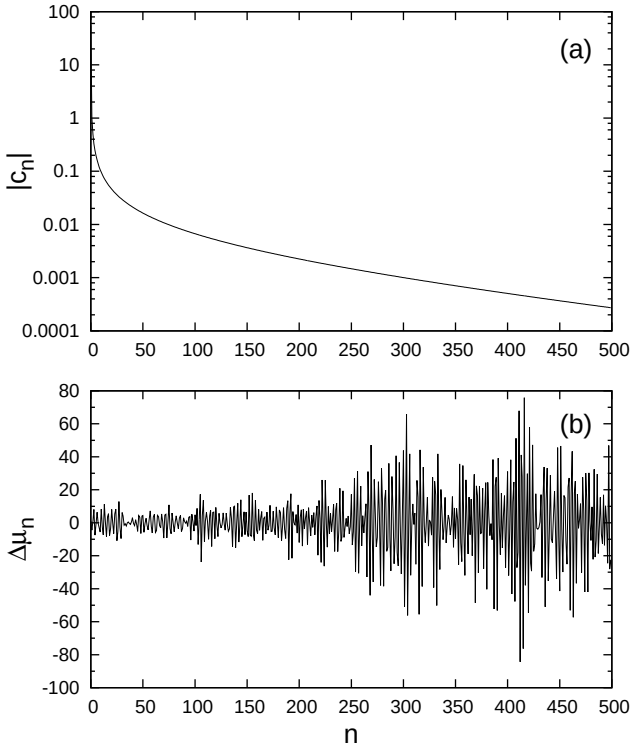

FIG. 6. $|c_n|$ (Eq. 13) and $\Delta\mu_n$ (Eq. 14) with respect to the moment index. The results are obtained for vacancy hopping, with the $15 \times 15 \times 15$ cell and 20000 random vectors.

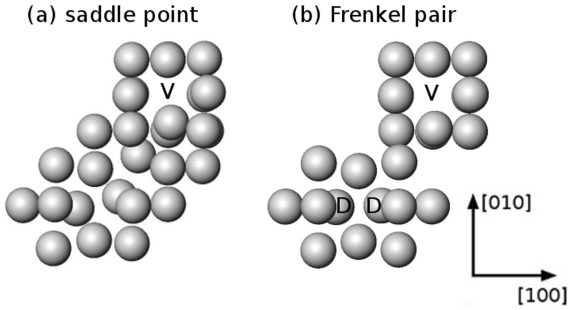

FIG. 7. Configurations of (a) the saddle point during the Frenkel pair formation and (b) the final Frenkel pair. For clarity, only non-FCC atoms are shown (as determined by a common-neighbor analysis). New vacancies are marked with "V", and the two atoms that form the dumbbell interstitial defect are marked with "D".

more moments are required to converge the prefactors for the Frenkel pair formation. The complexity of the DOS difference between the minimum and the saddle point is therefore the determining factor that controls the required order of the Chebyshev expansion.

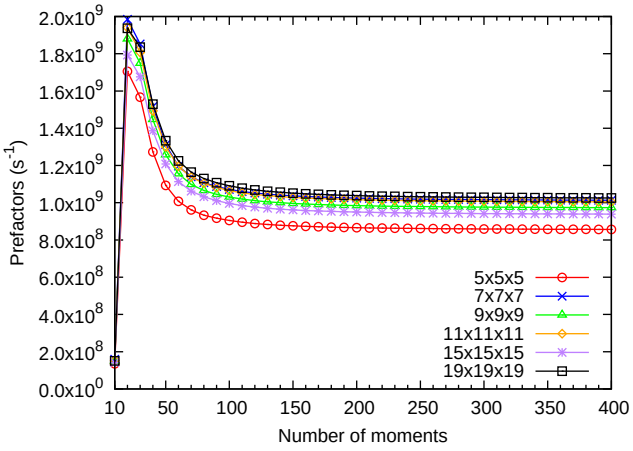

FIG. 8. Convergence of the prefactors for Frenkel pair formation (using 40000 random vectors) with respect to the number of moments, for five different cell sizes. For the $15 \times 15 \times 15$ cell, the KPM results converge to within 20%, 10%, 5%, and 1% error with 60, 80, 110, and 220 moments, respectively.

V. DISCUSSION

The results above demonstrate that our method is able to accurately compute HTST prefectors for large systems at an affordable cost. To put this in perspective, consider a system where 1000 moments and 1000 random vectors are required for high accuracy results. As each moment requires the calculations of the interatomic forces at the minimum and saddle point for each random vector, on the order of 10^6 force calculations are required overall. This corresponds to the same computational effort as a few nanoseconds of molecular dynamics (MD) simulation on the whole system, a non-negligible investment, but still very competitive compared to what can be expected from MD-based sampling techniques such as thermodynamic integration[8]. Furthermore, calculation corresponding to different random vectors can be done totally independently. Our method is therefore trivial to parallelize and extremely scalable as parallelization can be carried out over both individual force calculation (using spatial decomposition techniques) and over random vector realizations. The wall-clock time required for the calculation of the prefactor could therefore in principle be reduced to the time needed for, in this example, 1000 parallel force calculations. Based on the examples discussed here, we estimate that the cross-over point where the KPM method requires less computing effort than the direct diagonalization to be between the 15x15x15 and the 19x19x19 cell, assuming 40000 random vectors, which provides an error at the level of a few percents. We emphasize again however that the KPM can be massively parallelized over the different random vector realization with no loss in efficiency, which enables faster wall-clock time solutions at even smaller cell sizes. Another strength of our method is that the level of accuracy is tunable.

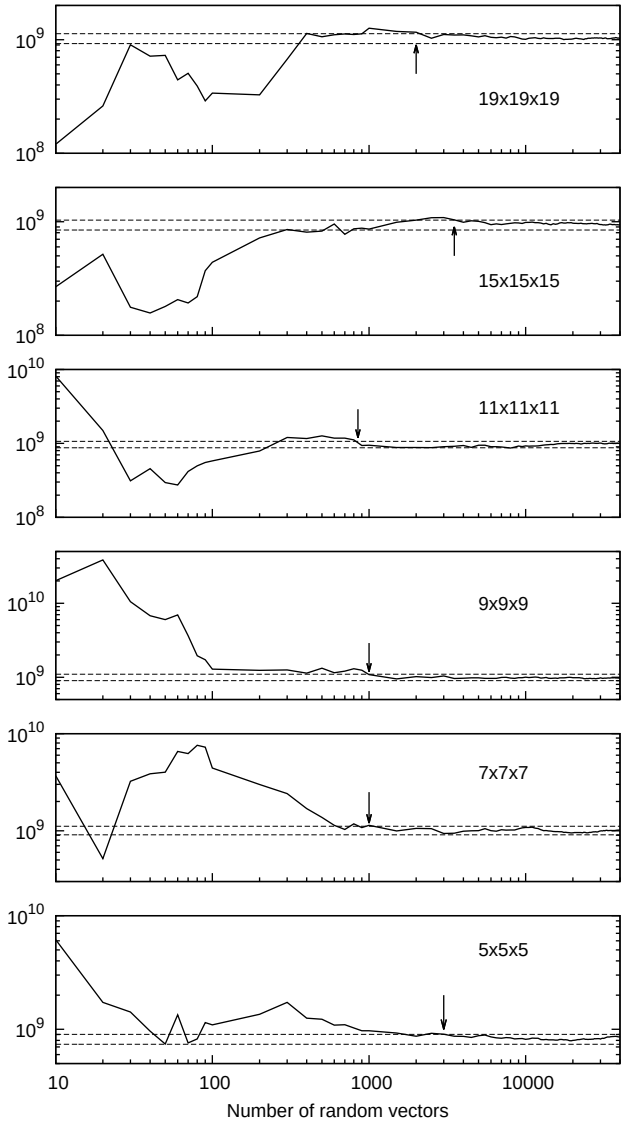

FIG. 9. Convergence of the prefactors (s^{-1}) for Frenkel pair formation (using 400 moments) with respect to the number of random vectors, for five different cell sizes. The dashed horizontal lines show the +/- 10% error windows from the benchmarks obtained by directly diagonalizing the Hessian matrices (except for the $19 \times 19 \times 19$ cell, where the benchmark is taken from the converged KPM results at 40000 random vectors). Arrows indicate convergence to within 10% error.

Based on the results given above, errors at the 50% levels can be achieved with only tens of random numbers and moments, which in turn costs only a few picoseconds worth of MD, a routine calculation even on systems containing millions of atoms. Finally, this method only requires interatomic forces, a quantity that is available from any MD code. Implementation of our method into existing codes is therefore straightforward.

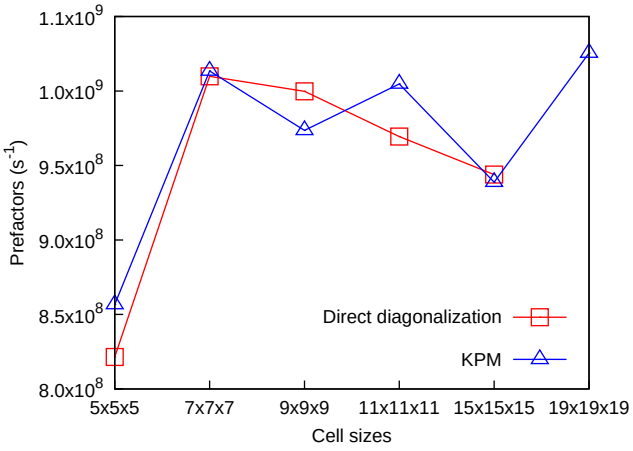

FIG. 10. Comparison of the prefactors of the Frenkel pair formation between the KPM and the direct diagonalization, for different cell sizes. The results are obtained with 400 moments and 40000 random vectors. The relative errors between the KPM and the direct diagonalization are 4.3%, 0.3%, 2.6%, 3.6% and 0.5% for the cell sizes from $5\times5\times5$ to $15\times15\times15$.

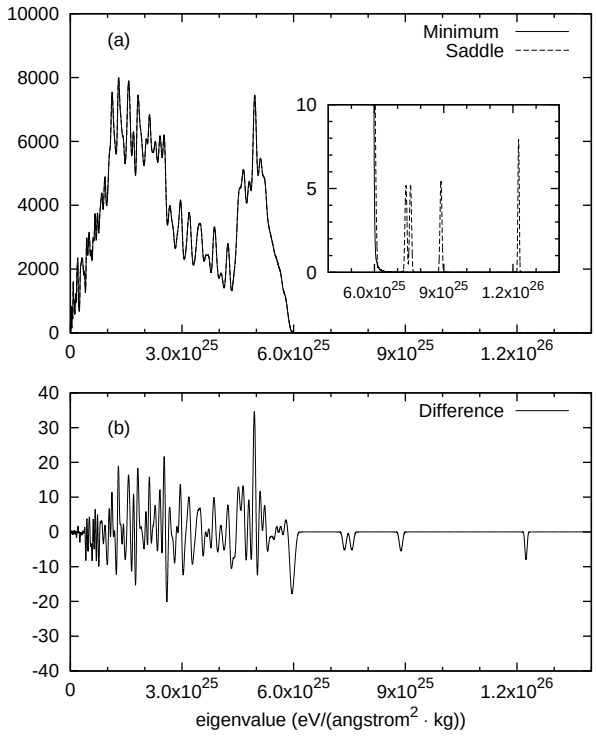

FIG. 11. (a) Comparison of the DOSs between the minimum and the saddle point for the vacancy hop. The results are obtained with a cell size of $15\times15\times15$, 400 moments, and 20000 random vectors. The inset magnifies the new peaks formed in the saddle point's DOS. (b) The difference of the DOS between the minimum and the saddle point.

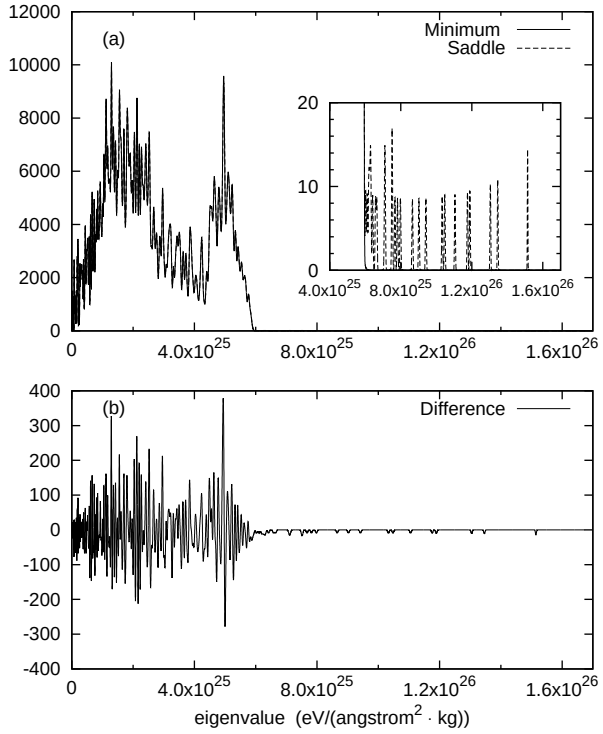

FIG. 12. (a) Comparison of the DOSs between the minimum and the saddle point for the Frenkel pair formation, for the cell size of $15\times15\times15$. The results are obained with 400 moments and 40000 random vectors. The inset magnifies the new peaks formed in the saddle point's DOS. (b) The difference of the DOS.

VI. CONCLUSIONS

Using the kernel polynomial method, we have developed an efficient technique for calculating the prefactors of transition rates in large systems within the HTST approximation. The method is based on the expansion of the vibrational DOS in terms of Chebyshev polynomials. As the order of the expansion is a user-tunable parameter, the method offers an adjustable balance between accuracy and computational cost. Using two prototypical processes, namely the diffusion of vacancies and the nucleation of Frenkel pairs in bulk silver, we demonstrated its high accuracy and efficiency. As the computational cost scales only linearly with system size for short-range potentials, it can be use to investigate important kinetic processes in very large systems.

ACKNOWLEDGMENTS

This work was supported by the United States Department of Energy (US DOE) Office of Science. Initial development of this method and final stages of the work were supported by the Office of Basic Energy Sciences, Division of Materials Sciences and Engineering. The middle

stage was supported by the Office of Advanced Scientific Computing Research. Los Alamos National Laboratory is operated by Los Alamos National Security, LLC, for the National Nuclear Security administration of the US DOE under contract DE-AC52-06NA25396.

* afv@lanl.gov
† danny_perez@lanl.gov

[1] J. P. Hirth and J. Lothe, *Theory of Dislocations*, 2nd ed. (Krieger Publishing Company, 1992).
[2] J. Friedel, *Dislocations*, 1st ed. (Pergamon press, 1964).
[3] G. H. Vineyard, J. Phys. Chem. Solids **3**, 121 (1957).
[4] P. A. Gordon, T. Neeraj, and M. J. Luton, Modelling Simul. Mater. Sci. Eng. **16**, 045006 (2008).
[5] T. Zhu, J. Li, A. Samanta, A. Leach, and K. Gall, Phys. Rev. Lett. **100**, 025502 (2008).
[6] S. Hara and J. Li, Phys. Rev. B **82**, 184114 (2010).
[7] L. Proville, D. Rodney, and M. Marinica, Nature Materials (2012).
[8] C. Jin, W. Ren, and Y. Xiang, Scripta Materialia **62**, 206 (2010).
[9] E. Vanden-Eijnden and M. Venturoli, J. Chem. Phys. **130**, 194103 (2009).
[10] S. Ryu, K. Kang, and W. Cai, Proceedings of the National Academy of Sciences **108**, 5174 (2011).
[11] A. Weiβe, G. Wellein, A. Alvermann, and H. Fehske, Review of Modern Physics **78**, 275 (2006).
[12] R. Silver and H. Röder, Int. J. Mod. Phys. C **5**, 935 (1994).
[13] L.-W. Wang, Phys. Rev. B **49**, 10154 (1994).
[14] L.-W. Wang and A. Zunger, Phys. Rev. Lett. **73**, 1039 (1994).
[15] A. F. Voter, J. D. Kress, and R. N. Silver, Phys. Rev. B **53**, 12700 (1996).
[16] S. Glasstone, K. J. Laidler, and H. Eyring, *The Theory of Rate Processes*, Vol. 19 (McGraw-Hill, New York, 1941).
[17] H. Eyring, J. Chem. Phys. **3**, 107 (1935).
[18] E. Wigner, Trans. Faraday Soc. **34**, 29 (1938).
[19] J. Horiuti, Bull. Chem. Soc. Jpn. **13**, 210 (1938).
[20] J. C. Keck, Discussions of the Faraday Society **33**, 173 (1962).
[21] J. C. Keck, Advances in Chemical Physics **13**, 85 (1967).
[22] G. Henkelman and H. Jónsson, The Journal of Chemical Physics **111**, 7010 (1999).
[23] G. Henkelman, B. P. Uberuaga, and H. Jónsson, The Journal of Chemical Physics **113**, 9910 (2000).
[24] R. N. Silver and H. Röder, Int. J. Mod. Phys. C **5**, 935 (1994).
[25] R. N. Silver, H. Röder, A. F. Voter, and D. J. Kress, J. of Comp. Phys. **124**, 115 (1996).
[26] J. Skilling, *Maximum Entropy and Bayesian Methods* (Kluwer Academic Publishers, Dordrecht, 1989) p. 455.
[27] J. C. Wheeler, Phys. Rev. A **9**, 825 (1974).
[28] D. A. Drabold and O. F. Sankey, Phys. Rev. Lett. **70**, 3631 (1993).
[29] P. Bratley, B. L. Fox, and L. E. Scharge, *A Guide to Simulation* (Springer-Verlag, New York, 1987).
[30] A. Hams and H. D. Raedt, Phys. Rev. E **62**, 4365 (2000).
[31] P. L. Williams, Y. Mishin, and J. C. Hamilton, Modelling Simul. Mater. Sci. Eng. **14**, 817 (2006).
[32] W. E, W. Ren, and E. Vanden-Eijnden, Phys. Rev. B **66**, 052301 (2002).
[33] C. Lanczos, J. Res. Nat. Bur. Standards **45**, 255 (1950).
[34] G. Box and M. Muller, Ann. Math. Statist. **29**, 610 (1958).
[35] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, 1992).
[36] R. Piessens, E. de Doncker-Kapenga, C. Überhuber, and D. Kahaner, *Quadpack: A Subroutine Package for Automatic Integration (Springer Series in Computational Mathematics*, 1st ed. (Springer, 1983).
[37] D. Jackson, PhD thesis (Georg-August-Universitt Gttingen) (1911).
[38] D. Jackson, Trans. Am. Math. Soc. **13**, 491 (1912).