Ridge-based bias potentials to accelerate molecular dynamics

Penghao Xiao, Juliana Duncan, Liang Zhang, and Graeme Henkelman (iD)

View Online   Export Citation   CrossMark

Ridge-based bias potentials to accelerate molecular dynamics

Penghao Xiao,[a] Juliana Duncan,[a] Liang Zhang, and Graeme Henkelman[b]
Department of Chemistry and the Institute for Computational and Engineering Sciences,
University of Texas at Austin, Austin, Texas 78712, USA

An effective way to accelerate rare events in molecular dynamics simulations is to apply a bias potential which destabilizes minima without biasing the transitions between stable states. This approach, called hyperdynamics, is limited by our ability to construct general bias potentials without having to understand the reaction mechanisms available to the system, *a priori*. Current bias potentials are typically constructed in terms of a metric which quantifies the distance that a trajectory deviates from the reactant state minimum. Such metrics include detection of negative curvatures of the potential, an energy increase, or deviations in bond lengths from the minimum. When one of these properties exceeds a critical value, the bias potentials are constructed to approach zero. A problem common to each of these schemes is that their effectiveness decreases rapidly with system size. We attribute this problem to a diminishing volume defined by the metrics around a reactant minimum as compared to the total volume of the reactant state basin. In this work, we mitigate the dimensionality scaling problem by constructing bias potentials that are based upon the distance to the boundary of the reactant basin. This distance is quantified in two ways: (i) by following the minimum mode direction to the reactant boundary and (ii) by training a machine learning algorithm to give an analytic expression for the boundary to which the distance can be calculated. Both of these ridge-based bias potentials are demonstrated to scale qualitatively better with dimensionality than the existing methods. We attribute this improvement to a greater filling fraction of the reactant state using the ridge-based bias potentials as compared to the standard potentials. © 2015 AIP Publishing LLC. [http://dx.doi.org/10.1063/1.4937393]

I. INTRODUCTION

Molecular dynamics (MD) is a ubiquitous tool used to capture the dynamic evolution of molecular systems. MD simulations use a finite difference approximation to integrate Newton's equations of motion with a time step on the order of femtoseconds. One of the main limitations of MD simulations is their inability to simulate molecular processes at time scales longer than ~1 ns due to the prohibitively large computational resources required. These slower events (often called "rare events") are important in many fields and include surface diffusion processes, surface catalysis, and grain boundary migration.

One successful method for accelerating MD simulations is Voter's hyperdynamics (HD) approach.[1] In HD, the potential energy surface of a chemical system is altered by adding a non-negative bias potential, $\Delta V(r)$.[2,3] Voter's derivation of HD shows that the reaction rate, calculated using the transition state theory (TST) approximation, is preserved when the bias potential goes to zero on the TST dividing surface, and the time evolution of a trajectory is accelerated according to

$$t_{\text{hyper}} = \sum_{i=1}^{n} \Delta t \; e^{\Delta V(r(t_i))/k_B T}. \tag{1}$$

Here, k_B is the Boltzmann's constant, T is the temperature, $r(t_i)$ is the $3N$ dimensional vector describing the position of the molecular system at time step i, N is the number of atoms, Δt is the integration time step on the biased potential, n is the number of time steps, and t_{hyper} is the time elapsed on the original PES. Then, the evolution of the trajectory on the biased potential maintains the correct state-to-state kinetics as the original PES with a computational acceleration or "boost" factor,

$$\text{boost} = \left\langle e^{\Delta V(r(t_i))/k_B T} \right\rangle_h = \frac{t_{\text{hyper}}}{n \Delta t}, \tag{2}$$

where the subscript h indicates an average over the trajectory on the biased hyper-potential.

One of the most challenging aspects of the HD approach is the construction of an effective bias potential that does not rely upon prior knowledge of the reaction mechanisms available to the chemical or materials system. Having a general bias potential that does not rely on a preconceived notion of how the system will react is important because it allows the HD method to reveal unexpected reaction mechanisms. An effective bias potential should also reduce barriers in the reactant state and go to zero at the reactant state boundary.

An important point that should be considered when constructing a bias potential is that it should not impede sampling of the reactant state. Poor sampling can result from regions of the bias potential that are not thermally accessible within the time scale of reactions on the hyper-potential. This consideration provides a conservative maximum bias set by

---

[a]P. Xiao and J. Duncan contributed equally to this work. P. X. implemented the HD-MMF method and J. D. implemented the HD-SVM method.
[b]Electronic mail: henkelman@utexas.edu

the lowest barrier to escape the reactant state. A large average boost factor cannot then be achieved by small regions of large boost; rather, a modest boost should be accumulated over as much of the reactant state as possible. The challenge lies in the fact that it is not easy to identify the dividing surface, which defines the boundary of the reactant state, in the high dimension systems of interest.

Several bias potentials have been suggested for the hyperdynamics method.[1–10] One of the primary limitations of current bias potentials is how their effectiveness scales with the dimensionality of the system. More specifically, most bias potentials give a large boost factor in low dimensional systems that falls off rapidly as the number of atoms in the system is increased. A point that we make in this paper is that these bias potentials are defined as being non-zero within some descriptor of the reactant minimum. If one thinks of this biased region as being similar to a hypersphere inside a larger hypercube, representing the true boundary of the reactant state, then the volume of the hypersphere drops rapidly with respect to the hypercube as the dimensionality increases. This argument motivates the construction of a bias potential that is made to go to zero near the boundary of the reactant state rather than be non-zero near the reactant minimum. The closer that the bias potential can be constructed as filling the reactant, the better the HD boost will scale with dimensionality. This argument is illustrated in Fig. 1.

Voter's original bias potential boosts the concave region near a minimum and regions where the dot product of the force and the min-mode direction is non-zero. The latter defines relatively narrow valleys connecting the minimum to saddle points. In a high dimensional space, however, the dot product between any pair of random vectors drops to zero and thus this bias leaves large areas between the concave region and the boundary of the reactant basin unbiased.

Another popular bias potential is the "bond-boost" method introduced by Miron and Fichthorn.[5] The bond-boost bias potential is based on how much the maximally stretched bond deviates from the equilibrium bond length. The method requires the user to select at what value the bias potential goes to zero, namely, the stretching threshold q. An appropriate value of q should be less than the maximum stretch of all

bonds at the transition states. The concept of boosting bonds can be related to the idea of boosting the concave region of the potential. Instead of using the min-mode direction as in Voter's bias, it uses the softest single bond stretching mode and instead of turning off the bias at inflection points, it uses an empirical value of the strain, which should ideally correspond to that of the true ridge. The advantage of the bond boost approach is that it naturally separates variables by using the maximum-stretched bond length. It is the summation over all degrees of freedom in the dot product operation of Voter's bias that leads to its poor dimensionality scaling. A drawback of the bond-boost bias is that it is turned off based upon a distance from the reactant minimum. In a high dimensional system, this boosted region becomes small even when the threshold, q, is set to that of the saddle points, where the bonds break. We should note that Voter and coworkers have recently developed a strategy to localize the bond-boost potential which improves the scaling for systems that are larger than considered in this work.[11]

A third kind of bias potential is the flat potential bias introduced by Steiner *et al.*[4,12] The flat bias turns off when the potential energy is above a specified level (typically up to the lowest saddle point) and sets the overall hyper-potential to the flat level otherwise. The flat bias can be thought of as a simplified version of the bond boost bias potential. If the total potential is approximated as a summation over pair-wise harmonic potentials along the bonds, then the bond boost potential shuts off when a maximum pair potential is reached whereas the flat potentials uses the sum over all the bonds. The flat-bias method does not scale as well as the bond boost because it couples together all degrees of freedom for determination of the bias, which reduces the volume of the boosted hypersphere with respect to the reactant basin with increasing dimension. Advantages of the flat bias are that it has no additional computational cost over regular MD and that it requires minimal effort to implement.

In this work, we present two bias potentials, both of which use the distance to the boundary of the reactant state (the ridge) in their construction. We show that this definition leads to an average boost factor that scales significantly better with dimensionality than the methods described above. The first ridge-based bias potential is determined on the fly by following the minimum curvature mode direction up to the ridge. In the second potential, the distance to the ridge is approximated with a support vector machine (SVM) algorithm trained with data from a high temperature MD trajectory. Both ridge-based bias potentials are turned off when the hyper-trajectory approaches the ridge.
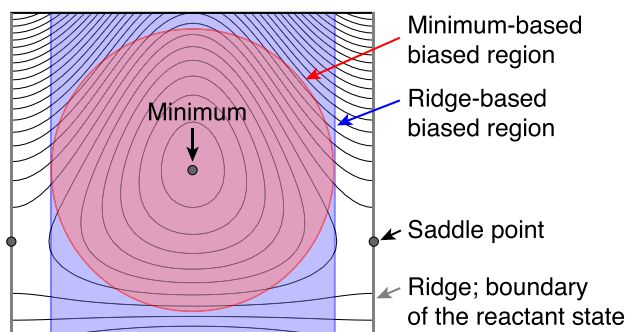

FIG. 1. A cartoon illustrating why minimum-based bias potentials can suffer from dimensionality scaling limitations more than ridge-based bias potentials. A reduction in the coverage of the reactant state becomes a severe problem as the dimensionality increases, in the same way that the ratio of a hypersphere to hypercube decreases rapidly with dimensionality.

II. DEFINITION OF THE RIDGE-BASED BIAS POTENTIALS

A. Ridge-based bias potential using minimum-mode following (MMF)

To determine the ridge-based bias potential using a MMF approach, a walker climbs up the potential energy surface $V(r)$ along the min-mode direction $\hat{\tau}$ until it escapes the reactant

basin. The MMF bias potential $\Delta V_b^{\text{MMF}}(r)$ is then defined as the energy difference between the escape point $r^{\ddagger}$ on the boundary of the reactant basin and the starting point of the walker, r, up to a maximum value of ΔV_b^{max},

$$\Delta V_b^{\text{MMF}}(r) = \begin{cases} V(r^{\ddagger}(r)) - V(r) & \text{if } V(r^{\ddagger}(r)) - V(r) < \Delta V_b^{\text{max}} \\ \Delta V_b^{\text{max}} & \text{if } V(r^{\ddagger}(r)) - V(r) \geq \Delta V_b^{\text{max}}. \end{cases}$$

(3)

Far from the ridge, the bias is the constant maximum value, ΔV_b^{max}; near the ridge, where $V(r^{\ddagger}(r)) - V(r) < \Delta V_b^{\text{max}}$, the hyper-potential, $V_h^{\text{MMF}}(r) \equiv V(r) + \Delta V_b^{\text{MMF}}(r)$, is equal to the energy of the escape point, $V(r^{\ddagger}(r))$. Since the energy of the min-mode walker increases monotonically, having a maximum bias limits the climbing distance. If the climber energy increases larger than ΔV_b^{max}, the climb is terminated and the maximum bias is applied. Using a short climbing distance not only lowers the computational cost for determining the bias but also avoids discontinuities of the bias where two or more ridges merge. Of course, the benefits of a small climbing distance need to be balanced with the goal of having a large maximum bias potential and boost factor.

In the fully boosted region, where $\Delta V_b^{\text{MMF}}(r) = \Delta V_b^{\text{max}}$, the bias potential is constant and the hyper-force is the same as the force on the unbiased potential. Outside the fully boosted region, near the ridge, the hyper-force is determined by the corresponding escape position on the potential ridge, $r^{\ddagger}(r)$, as calculated by the MMF climber. In fact, all points along the climb trajectory share the same escape point and therefore the same hyper-force. Thus, the component of the hyper-force along the climb trajectory, i.e., the minimum-mode direction, is zero,

$$\nabla V_h^{\text{MMF}}(r) = \begin{cases} \nabla V(r^{\ddagger})\mathbf{J}_{r^{\ddagger}}(r) & \text{if } V(r^{\ddagger}(r)) - V(r) < \Delta V_b^{\text{max}} \\ \nabla V(r) & \text{if } V(r^{\ddagger}(r)) - V(r) \geq \Delta V_b^{\text{max}}. \end{cases}$$

(4)

The hyper-force perpendicular to the climbing path is related to the perpendicular force at the escape point $r^{\ddagger}(r)$. Specifically, the perpendicular force should be rotated according to the change in direction of the minimum-mode direction along the climbing path and scaled in magnitude according to the change in a volume element or Jacobian along the climbing path, $\mathbf{J}_{r^{\ddagger}}(r)$. When the maximum bias is small enough, the $V(r^{\ddagger}(r)) - V(r) < \Delta V_b^{\text{max}}$ region is near the ridge and $\mathbf{J}_{r^{\ddagger}}(r)$ in Eq. (4) can be approximated by a rotation matrix which aligns $\tau(r^{\ddagger})$ with $\tau(r)$ and thus makes $\nabla V_h^{\text{MMF}}(r)$ perpendicular to $\tau(r)$. In practice, we find that $\mathbf{J}_{r^{\ddagger}}(r)$ can be approximated as the identity matrix with little effect on the accuracy of calculated escape rates. In this way, the hyper-force is simply $\nabla V(r^{\ddagger})$ as calculated by the min-mode walker.

A schematic picture of our ridge-based MMF bias potential strategy is shown in Fig. 2. A trajectory spends most of its time in the maximally boosted region. Once it escapes to the flat region where $V(r^{\ddagger}(r)) - V(r) < \Delta V_b^{\text{max}}$, it is unlikely to return to the reactant state. This bias is different from Voter's original bias potential which uses a force projection along the min-mode direction. The ridge-based bias potential is zero at
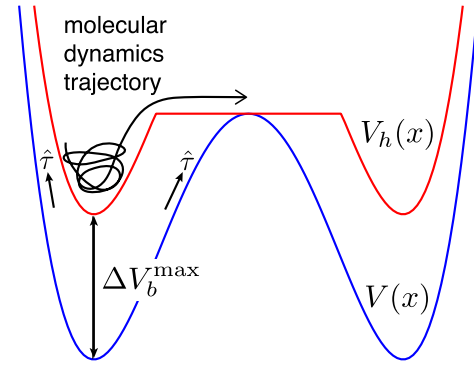

FIG. 2. Illustration of the ridge-based bias potential.

the true ridge, while Voter's goes to zero at a local definition of the ridge, where the dot product of the force and the negative mode is zero. Importantly, the local ridge can be qualitatively different from the true ridge.[13] Moreover, our ridge-based bias flattens the potential energy surface, does not introduce any new barriers, and thus facilitates efficiently sampling in the reactant region. Most importantly, the ridge-based bias potential scales well with dimensionality, as we will show subsequently.

To minimize the cost of calculating the MMF bias potential, several computational tricks are employed.

1. If a trajectory is in the maximal bias region, ΔV_b is calculated only every five MD steps to check if it equals ΔV_b^{max}; for the other steps, we take $\Delta V_b = \Delta V_b^{\text{max}}$ and $F_{\text{hyper}} = \nabla V(r)$ without incurring any extra computational cost.

2. When the previous check shows $\Delta V_b = \Delta V_b^{\text{max}}$, the numerical setting for the next ΔV_b check can be very liberal. We increase the step size for min-mode climbing to 0.4 Å (from 0.1 Å) and the min-mode rotational tolerance to 3° (from 1°).

3. In concave regions, where there is no negative curvature mode, we take $\Delta V_b = \Delta V_b^{\text{max}}$ without performing any min-mode climbing.

4. Only when the condition that $\Delta V_b < \Delta V_b^{\text{max}}$ is detected, ΔV_b is updated at every step using tight convergence settings, where the climbing step size is 0.1 Å and the dimer rotation tolerance is 1°.

One problem with the MMF method is that it may not detect a ridge in a part of the potential where there are two negative modes. At low temperature, when rare events occur near first-order saddle points, this is not a significant concern. At high temperature, however, trajectories may leave the reactant basin where there is at least one negative mode parallel to the ridge as well as the one across it. If a parallel mode has the most negative curvature, the MMF walker will climb parallel to the ridge and not detect it. To test for and to correct this problem, we propose that MMF walkers can be used to follow all of the negative modes. The bias potential will then be the minimum of that from each walker. In the following example, we found only a 10% increase in the rate at the highest temperature considered (600 K in the Al/Al(100) system) when all negative modes were followed and so all

results are based upon following just the lowest curvature mode.

A second issue related to the accuracy of the MMF method is that the flat region introduced by the bias potential at the ridge will increase recrossing effects and decrease the Kramers transmission coefficient. In a recent paper by Huang, Perez, and Voter,[14] where an efficient bias potential based on predefined minimum energy paths is defined, it is shown that a buffer of unbiased potential near the ridge can be introduced to recover the recrossing effects of MD on the original potential. The same strategy can be applied here by making a minor change to the bias potential,

$$\Delta V_b^{\text{MMF}}(r) = \max\{\Delta V_b^{\text{MMF}}(r) - \Delta V_b^{\text{safe}}, 0\} \quad (5)$$

$$\text{if } \Delta V_b^{\text{MMF}}(r) = 0 \text{ then } \nabla V_h^{\text{MMF}}(r) = \nabla V(r), \quad (6)$$

where ΔV_b^{safe} is the height of the "safety" zone at the ridge which allows for proper recrossing effects. The following examples are implemented with $\Delta V_b^{\text{safe}} = 0$ since such dynamical corrections were not found to be important in our test cases.

B. Ridge-based bias potential using machine learning

The second way that we construct a ridge-based bias potential is by first characterizing the boundary of the reactant state prior to running hyperdynamics. We do this by training a machine learning algorithm to classify structures that are known to belong to the reactant state. An appropriate tool for this classification problem is the one-class support vector machine (OCSVM).[15,16] The OCSVM requires a set of data points from a distribution of a single class that describes the given distribution, i.e., from the reactant state. The algorithm produces a decision function, $f(r)$, which is positive within the reactant state and negative outside. In this way, the OCSVM finds a hypersurface where the decision function is equal to zero. The training set for the OCSVM is generated using high temperature MD that is constrained to the reactant state. During the MD trajectory, the system is periodically minimized to detect if the system escapes the reactant state. If the system is found to escape the reactant state, points belonging to product states are discarded and the trajectory is restarted within the reactant state. Details of the OCSVM can be found in the supplementary material.[17]

In the language of machine learning, we use the Gaussian radial basis function kernel to describe the distance between points in feature space,

$$K(x_i, x_j) = e^{-\frac{|x_i - x_j|^2}{2\gamma^2}}, \quad (7)$$

where γ controls the width of the Gaussian basis functions. The appropriate choice of γ is non-trivial, and in general, parameter selection plays an important role in the capability of the OCSVM to model complex distributions. Small values of γ lead to over-fitting and many small Gaussian functions around the training points. Large values of γ lead to an over-generalization of the distribution which can fail to capture the details of the class boundary. Some guidelines for the selection of γ are described later.

One challenge that we faced when generating a SVM bias potential is that while the OCSVM algorithm is designed to classify points in the reactant state, it does not provide information as to how close the point is to the boundary of the reactant state. Typically, OCSVM decision functions have maximum (absolute) values within a distance γ of the class boundary and small values well inside or outside of the reactant state. The non-monotonic nature of $f(r)$ makes it unsuitable to use directly as a bias potential. Instead, we use a different metric: the distance to the closest point on the OCSVM boundary, $r_{\text{surf}}(r)$,

$$d(r) = |r_{\text{surf}}(r) - r|. \quad (8)$$

Given a boundary generated from the OCSVM, the bias potential $\Delta V_b^{\text{SVM}}(r)$ is defined by

$$\Delta V_b^{\text{SVM}}(r) = \begin{cases} \Delta V_b^{\max}\left(\dfrac{d(r)^2}{2w^2} - \dfrac{d(r)^3}{3w^3}\right) & \text{if } d(r) < w \text{ and } f(r) > 0 \\[2mm] \Delta V_b^{\max} & \text{if } d(r) \geq w \text{ and } f(r) > 0 \\[2mm] 0 & \text{if } f(r) \leq 0 \end{cases}, \quad (9)$$

where w is a cutoff distance for the constant region of the bias potential, as described previously for the MMF bias potential. The gradient of the SVM potential is

$$\nabla V_h^{\text{SVM}}(r) = \begin{cases} \nabla V(r) + \Delta V_b^{\max}\left(\dfrac{d(r)}{w^2} - \dfrac{d(r)^2}{w^3}\right)\hat{d}(r) & \text{if } d(r) < w \text{ and } f(r) > 0 \\[2mm] \nabla V(r) & \text{if } d(r) \geq w \text{ and } f(r) > 0 \\[2mm] \nabla V(r) & \text{if } f(r) \leq 0 \end{cases}, \quad (10)$$

where $\hat{d}(r)$ is a unit vector connecting r and $r_{\text{surf}}(r)$. Figure 3 provides a schematic of how the SVM bias potential is constructed.

For the HD-SVM algorithm, we need to know the closest distance between r and the class boundary. The algorithm for doing this is illustrated in Fig. 4 with additional details
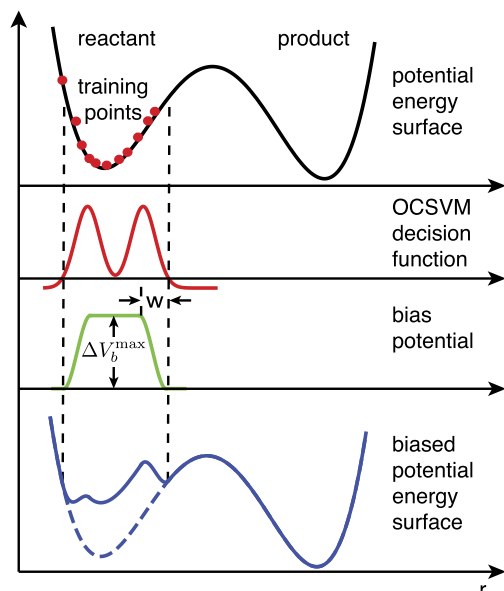
FIG. 3. Illustration of how the SVM bias potential is constructed.

provided in the supplementary material.[17] Briefly, we first find the nearest support vector and use ∇f with Newton's method to find a collinear point between r and the support vector which lies on the $f = 0$ surface. Second, we again use ∇f to minimize the distance to r while remaining on the $f = 0$ surface. These two steps are repeated iteratively until the closest point on the boundary is found to be within a specified tolerance. The resulting distance from r to the boundary, $d(r)$, is used to determine the bias potential from Eq. (9).

In Sec. III, we first show how the ridge-based bias potentials are constructed for a two dimensional system and demonstrate their accuracy for rate calculations. Second, we compare the dimensionality scaling of the HD boost factor of our ridge-based bias potentials in comparison with the existing bias potentials mentioned in the Introduction. Finally, we calculate the overall computational speed-up of running accelerated MD, including the cost associated with the ridge-based bias potentials, and show that the net gains are significant.
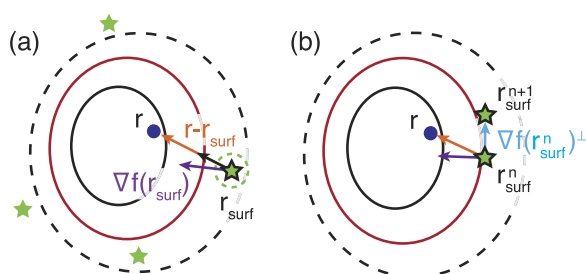

FIG. 4. (a) A method to find the closest point on the class boundary (red line). Given an initial point, r, the closest support vector (green star) is found. Newton's method, based upon the gradient of the decision function, ∇f, is used to determine the collinear point on the decision surface. (b) Newton's method is also used to move perpendicular to the hypersurface to find the closest point to r on the class boundary.

III. RESULTS FOR A TWO-DIMENSIONAL POTENTIAL

Our first simple test is for the "Voter97" two dimensional potential of the form[2]

$$V(x, y) = \cos(2\pi x)(1 + 4y) + \frac{1}{2}(2\pi y)^2. \quad (11)$$

The potential is harmonic in the y-direction and periodic in the x-direction. Each local minimum has two saddle points leading to symmetric minima towards $-x$ and $+x$ with identical energy barriers of 2.0.

A. Minimum-mode following bias potential

The hyper-potential produced by min-mode following is plotted in Fig. 5 together with the original potential. The bias potential decreases the barrier and smooths the potential within the minimum basin; no extra roughness is introduced in the hyper-potential. Figure 6 shows that the rates to escape the reactant basin at various temperatures are accurate as compared to direct MD simulations. The average boost factor $\langle e^{\beta \Delta V_b} \rangle$ scales exponentially with $1/k_B T$. The slope in Fig. 6 is equal to $\Delta V_b^{max} = 1.0$, which corresponds to the ideal case (grey dashed line) in which the trajectories remain in the fully boosted region before escaping the reactant state. These results indicate that the MMF ridge-based bias potential is doing well as can be expected for this model system.

B. Support vector machine bias potential

In order to train the OCSVM, a set of 3000 points was generated by running a MD trajectory constrained to the reactant state at $k_B T = 0.35$. As mentioned previously, the width of the OCSVM Gaussian kernel, γ, is required as an

(a) potential, $V(r)$ (b) bias potential, $\Delta V_b(r)$

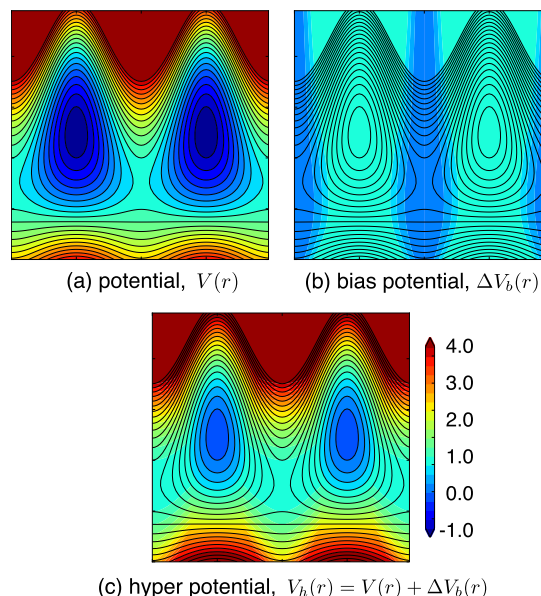(c) hyper potential, $V_h(r) = V(r) + \Delta V_b(r)$

FIG. 5. Illustration of the MMF ridge-based bias as applied to (a) the Voter97 potential. (b) The MMF bias potential and (c) the resulting hyper-potential are shown by the color scale on top of the black contour lines of the original potential. The maximum bias, ΔV_b^{max}, is set to be 1.0, which is half of the barrier height.
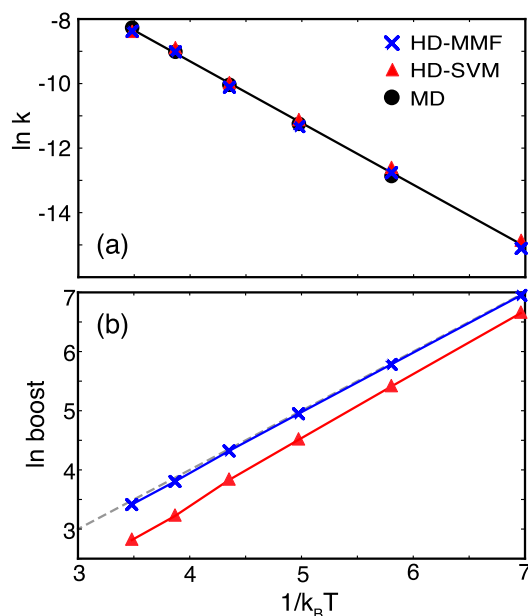
FIG. 6. (a) Comparison of the escape rate from the Voter97 potential minimum calculated using MD and HD with the MMF and SVM ridge-based bias potentials. (b) The average boost factor for the HD calculations; the ideal case is indicated by the dashed grey line.

input to the training process. Figure 7 illustrates a challenge associated with the selection of γ. When γ is too small, as in Fig. 7(b), over-fitting leads to islands (where $f < 0$) in the reactant state. When γ is too large, as in Fig. 7(d), the OCSVM over-generalizes the data set and provides insufficient resolution of the reactant state boundary. Figure 7(c) represents a suitable choice of γ. To test the HD performance as a function of γ, the maximum theoretical boost is determined at the temperature for which the training set was generated. For this test, we ran dynamics on the potential energy surface and assigned a maximum boost of $\exp(\Delta V_b^{max}/k_B T)$, where $\Delta V_b^{max} = 1.0$ and $k_B T = 0.35$ when the trajectory was inside the $f = 0$ surface and zero otherwise. The results, in Fig. 7(a),
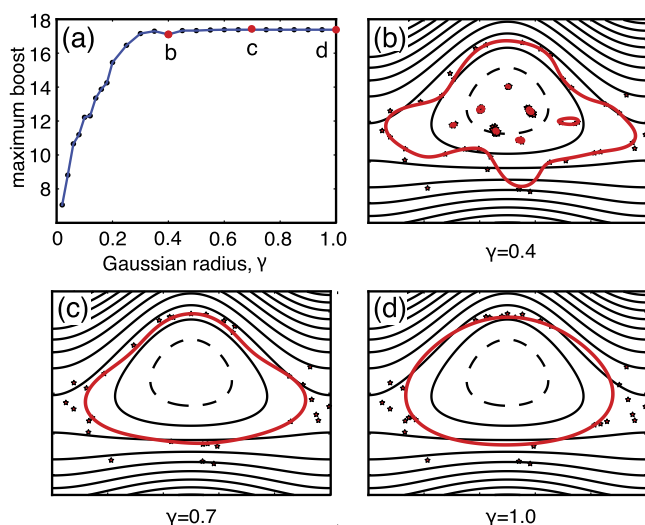
show that this test is not able to distinguish between functions f that over-fit, over-generalize, or optimally classify the data set. However, this sort of test will be useful for finding how γ should scale with dimensionality.

Figure 6(a) shows that the HD-SVM calculated rates to escape the reactant basin are accurate as compared to direct MD simulations over a wide range of temperatures. Values of γ and w were set to 0.7 and 0.15, respectively. The average boost factors, in Fig. 6(b), are somewhat lower than the ideal case indicating that the SVM method is not capturing the reactant state boundary quite as well as the MMF method.

IV. EFFICIENCY SCALING ON A HIGHER DIMENSIONAL POTENTIAL

We now show how the ridge-based bias potential methods scale with dimensionality and compare them with other approaches. Our test system is that of an Al adatom on the Al(100) surface. The atomic interactions are described by an embedded atom method potential.[18] To construct a low-dimensional version of this system, all atoms except the adatom are fixed to generate a three dimensional potential energy surface. The dimensionality is increased by relaxing the first, second, and third nearest neighbor shells around the adatom. These four relaxed shell structures, with 3, 15, 54, and 102 degrees of freedom, are shown in Fig. 8. In the 102D case, the two atom exchange reaction has the lowest diffusion barrier of 0.32 eV and the hop mechanism has a barrier of 0.38 eV; the hop is favored when fewer degrees of freedom are relaxed.

A. Parameter selection

Before comparing the efficiency of various HD bias potentials, the issue of parameter selection for the OCSVM needs to be addressed. This is a hard problem in high
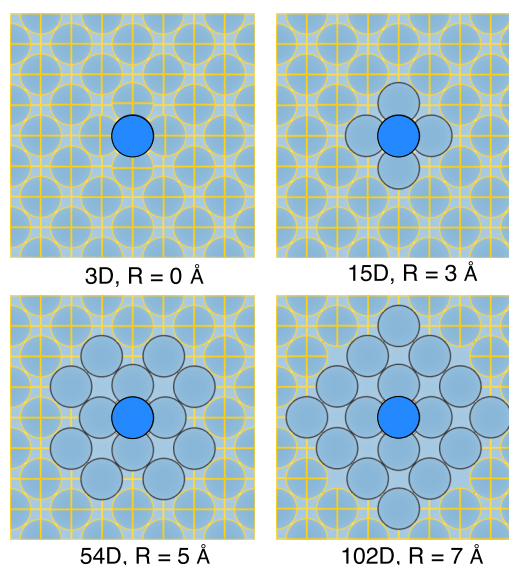

FIG. 7. (a) Plot showing how the maximum possible boost varies with the width of the OCSVM kernel, γ. The class boundary at $f = 0$ for γ equal to (b) 0.4, (c) 0.7, and (d) 1.0.

FIG. 8. Al adatom systems with increasing numbers of degrees of freedom based upon fixing atoms beyond a radius R from the adatom (dark blue).

dimensional systems where the surface cannot be easily visualized as in Fig. 7. Instead, we can first determine how the Gaussian radius γ should scale with dimensionality. Using the same methodology as in Fig. 7(a), we calculate how the maximum possible boost factor varies for HD-SVM trajectories calculated as a function of the Gaussian kernel width. A fixed training set size of 20 000 points was generated from a MD trajectory at 400 K; the HD-SVM calculation was also done at 400 K using $\Delta V_b^{max} = 0.2$. In Fig. 9(a), it is clear that for small γ, the boost increases rapidly with γ, indicating that the reactant region is over-fit around the training points, as in Fig. 7(b). At large γ, on the other hand, where the maximum boost saturates, we expect to be in the under-fitting regime. The values of γ for which the boost first saturates (open circles) are between the undesirable extremes. This critical γ value that lies between the under- and over-fitting regimes scales with the square-root of the dimensionality, as shown in Fig. 9(b). This result is expected because it is how the magnitude of a multi-dimensional vector, with equivalent components in each dimension, increases with dimensionality. From this result, we conclude that the optimal value of γ can be determined for one system (numerically) and then scaled according to the dimensionality of similar systems. Using the same logic, values of w were also scaled with the square-root of the number of degrees of freedom.

The system for which the optimal HD-SVM parameters were determined was the 102D Al adatom system. As was seen in Fig. 7 for the 2D example, the critical value of γ may
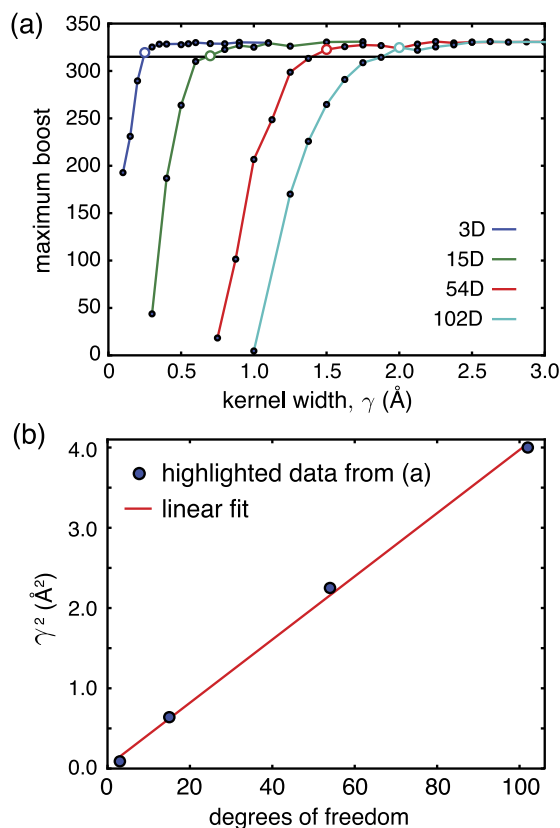
not be optimal. With these results in mind, we choose a value close to the critical value of γ as well as one twice as large. In Fig. 10, the boost factors obtained with a hyper-potential fit for these two values of γ are plotted as a function of the number of OCSVM training points. The HD tests were done at 350 K using the smallest value of $w = 0.4$ Å that gave accurate escape rates. For both γ values, the boost factor was found to drop with insufficient training data, although the larger value of $\gamma = 4.6$ Å required fewer training points to reach a saturated boost. In principle, there could be more tuning of these parameters, particularly with respect to determination of a maximum γ that is acceptable without over-generalization. For the purposes of this study, however, we were content with a SVM biased potential based upon 20 000 training points generated at 400 K, fit using a kernel width of $\gamma = 4.6$ Å, and a switching distance $w = 0.4$ Å, where these two distance parameters were scaled by the square root of the number of degrees of freedom for the smaller systems (Fig. 11).

B. Performance of the ridge-based bias potential

The key question being asked here is how the ridge-based bias potentials scale with dimensionality as compared to other bias potentials. For this test, we compare the MMF and SVM potentials with Voter's original bias potential, the bond boost method, and the flat bias. For the bond-boost method, two values of the bond-stretch threshold are compared to that of $q = 0.2$ (BB-q0.2) and $q = 0.3$ (BB-q0.3). For the Al adatom system, the saddle point for the lowest energy exchange mechanism has a maximum bond stretch of 37% at the saddle point, which indicates that the value of $q = 0.2$ is probably safe and $q = 0.3$ is slightly aggressive. The maximum bias for the potentials was set to 0.2 eV, except for the flat bias which was set at 0.3 eV above the reactant minimum, just below the lowest energy saddle. The average boost factors were calculated by running sufficiently long HD trajectories at 400 K in the reactant state. In the case of the Voter bias, Monte Carlo sampling of the hyper-potential was used instead of MD to avoid calculation of the hyper-force.

In the 3D case, the boost factors from all of the bias potentials are on the order of 100. When the number of
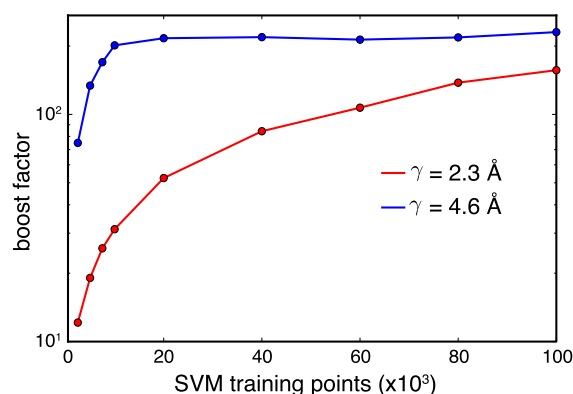

FIG. 9. (a) The maximum possible boost for SVM-classified reactant states as a function of the kernel width, γ. (b) The optimal kernel width squared scales linearly with the number of dimensions.

FIG. 10. The boost factor produced for various training set sizes for two values of γ.
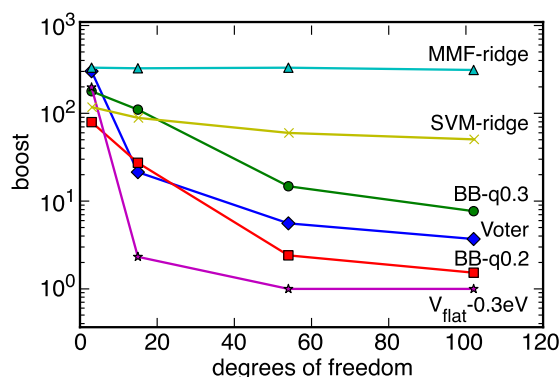
FIG. 11. Average boost factor for a set of different bias potentials calculated as a function of the number of degrees of freedom in the system.

dimensions increases to 15, where only five atoms are free to move, the non-ridge-based bias methods lose more than 50% of their boost factor. At 102 dimensions, which is characteristic of a realistic system, their boost factors have dropped by at least an order of magnitude. In contrast, the ridge based methods largely maintain their boost factors as the number of degrees of freedom increases.

The accuracy and performance of the ridge-based bias potentials at various temperatures are summarized in Table I. The net acceleration of each method is calculated as the average boost divided by the average number of force evaluations required per HD time step. Clearly the ridge-based bias potentials are more expensive to compute as compared to other bias potentials so that a sufficient boost factor is required to make them more efficient than running MD. The overall performance of the methods will be a function of temperature both because the boost increases rapidly with lowering temperature and because the ridge-based bias potentials require less computational work to evaluate when they are far from the ridge and the bias potentials are constant. On this fully boosted region, only occasional checks are required to see if the trajectory is approaching a ridge.

In the two ridge-based methods, the computational overhead occurs at different points. In the HD-MMF method, there is no initial overhead but there is extra work at each HD step. At 400 K, the HD-MMF method requires an average of 17 force evaluations per HD step with a boost factor of 325 resulting in an overall acceleration of 19. In the HD-SVM method, the overhead is required initially to generate OCSVM training points. Taking our choice of 20 000 training points, each is minimized using approximately 100 force evaluations, giving a fixed overhead of 2×10^6 force evaluations. For a conservative estimation of the acceleration of the HD-SVM method, we can assume that the bias potential will only be used for a single escape of the reactant state. At 400 K, a couple of million HD iterations are required to see a transition so that the overhead is roughly a factor of two. With a boost factor of 47, the overall acceleration is approximately 24. In summary, both the HD-MMF and HD-SVM methods provide accurate escape times for this test system with acceleration factors of about 20 as compared to MD at 400 K. This is twice the acceleration of the bond boost method using the most aggressive bond stretch parameter of 0.3.

V. CONCLUSION

Two forms of ridge-based bias potentials are described: one based upon a direct on-the-fly minimum-mode following trajectory to the ridge, and the other based upon a distance to a machine-learned boundary of the reactant state. Both ridge-based bias potentials are found to scale better with dimensionality than existing bias potentials. We attribute this difference to the ridge-based definition of the reactant basin as a distance to the boundary of the reactant state rather than to its minimum.

TABLE I. Rates ($\times 10^{-6}$ fs^{-1}), boost factors, and efficiency of the ridge-based bias potentials as calculated on the Al adatom system with 102 degrees of freedom.

Temperature (K)	600	500	400	300
TST rate	20(1)	5.1(7)	0.6(1)	0.03(1)
HD-MMF rate	16(1)	5(1)	0.59(7)	0.028(2)
Boost	45	99	325	2271
⟨Force calls/step⟩	39	34	17	8
Acceleration	1.2	2.9	19	284
HD-SVM rate	17(2)	4.1(5)	0.55(7)	0.013(2)
Boost	1.7	4.7	47	1739

[1]A. F. Voter, F. Montalenti, and T. C. Germann, Annu. Rev. Mater. Res. **32**, 321 (2002).
[2]A. F. Voter, J. Chem. Phys. **106**, 4665 (1997).
[3]A. F. Voter, Phys. Rev. Lett. **78**, 3908 (1997).
[4]M. M. Steiner, P. A. Genilloud, and J. W. Wilkins, Phys. Rev. B **57**, 10236 (1998).
[5]R. A. Miron and K. A. Fichthorn, J. Chem. Phys. **119**, 6210 (2003).
[6]S. Hara and J. Li, Phys. Rev. B **82**, 184114 (2010).
[7]X. Gong and J. Wilkins, Phys. Rev. B **59**, 54 (1999).
[8]Q. F. Fang and R. Wang, Phys. Rev. B **62**, 9317 (2000).
[9]S. Pal and K. A. Fichthorn, Chem. Eng. J. **74**, 77 (1999).
[10]J. Wang, S. Pal, and K. A. Fichthorn, Phys. Rev. B **63**, 085403 (2001).
[11]S. Y. Kim, D. Perez, and A. F. Voter, J. Chem. Phys. **139**, 144110 (2013).
[12]P. Tiwary and A. van de Walle, Phys. Rev. B **84**, 100301R (2011).
[13]P. Xiao, Q. Wu, and G. Henkelman, J. Chem. Phys. **141**, 164111 (2014).
[14]C. Huang, D. Perez, and A. F. Voter, J. Chem. Phys. **143**, 074113 (2015).
[15]D. Tax and R. Duin, Pattern Recognit. Lett. **20**, 1191 (1999).
[16]B. Schölkopf, J. Platt, J. Shawe-Taylor, A. Smola, and R. Williamson, Neural Comput. **13**, 1443 (2001).
[17]See supplementary material at http://dx.doi.org/10.1063/1.4937393 for details of the ridge-based bias potential using (I) minimum mode following and (II) machine learning.
[18]A. F. Voter and S. P. Chen, Mater. Res. Soc. Symp. Proc. **82**, 175 (1986).