Transition path sampling of rare events by shooting from the top

Hendrik Jung, Kei-ichi Okazaki, and Gerhard Hummer

Citation: The Journal of Chemical Physics **147**, 152716 (2017);
View online: https://doi.org/10.1063/1.4997378
View Table of Contents: http://aip.scitation.org/toc/jcp/147/15
Published by the American Institute of Physics

Transition path sampling of rare events by shooting from the top

Hendrik Jung,[1,2] Kei-ichi Okazaki,[1,3] and Gerhard Hummer[1,4,a)]
[1]*Department of Theoretical Biophysics, Max Planck Institute of Biophysics, Max-von-Laue Str. 3, 60438 Frankfurt am Main, Germany*
[2]*Department of Physics, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany*
[3]*Department of Theoretical and Computational Molecular Science, Institute for Molecular Science, National Institutes of Natural Sciences, Okazaki, Aichi 444-8585, Japan*
[4]*Institute of Biophysics, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany*

(Received 23 May 2017; accepted 24 July 2017; published online 14 August 2017)

Transition path sampling is a powerful tool in the study of rare events. Shooting trial trajectories from configurations along existing transition paths proved particularly efficient in the sampling of reactive trajectories. However, most shooting attempts tend not to result in transition paths, in particular in cases where the transition dynamics has diffusive character. To overcome the resulting efficiency problem, we developed an algorithm for "shooting from the top." We first define a shooting range through which all paths have to pass and then shoot off trial trajectories only from within this range. For a well chosen shooting range, nearly every shot is successful, resulting in an accepted transition path. To deal with multiple mechanisms, weighted shooting ranges can be used. To cope with the problem of unsuitably placed shooting ranges, we developed an algorithm that iteratively improves the location of the shooting range. The transition path sampling procedure is illustrated for models of diffusive and Langevin dynamics. The method should be particularly useful in cases where the transition paths are long so that only relatively few shots are possible, yet reasonable order parameters are known. [http://dx.doi.org/10.1063/1.4997378]

I. INTRODUCTION

Dynamics in complex systems is often dominated by rare events, from the nucleation of crystals to the folding of proteins.[1,2] The typical waiting time for the events of interest is much longer than the event itself. Therefore, even if it is technically possible to propagate large systems long enough in simulation time to witness a rare event, this direct sampling is not efficient when the relevant transition makes up only a tiny fraction of the trajectory. Transition path sampling (TPS) tries to overcome the disparity of time scales through importance sampling in trajectory space.[1,3] A Markov process in trajectory space is constructed whose target stationary distribution is the distribution of transition paths that would be obtained in long equilibrium runs.

Within a wide spectrum of sampling strategies, transition path (TP) shooting emerged as particularly powerful sampling strategy.[4] A point along an existing TP is chosen at random, its velocities are perturbed, and then trajectory segments are propagated forward and backward in time starting from the perturbed phase point. If the two trajectory segments end up in different metastable states, the new path is accepted, after reversing time on the backward segment. Otherwise, it is rejected. This procedure is then applied recursively.

For TP shooting to be efficient, it is important that paths are short and decorrelate after a few iterations. This requires a reasonably high degree of acceptance of trial paths and good connectivity of the TP ensemble under the iterative TPS procedure. Unfortunately, in many interesting problems, the probability of acceptance is low for almost any newly generated path. The probability of creating a TP starting from a phase point $x = (\boldsymbol{r}, \boldsymbol{v})$ is $p(\mathrm{TP}|x) = \phi_A(x)\phi_B(\underline{x}) + \phi_B(x)\phi_A(\underline{x})$,[5] where $\phi_A(x)$ and $\phi_B(x) = 1 - \phi_A(x)$ are the committor functions, defined as the probability of arriving first at states A and B, respectively. The underline in $\underline{x} \equiv (\boldsymbol{r}, -\boldsymbol{v})$ indicates sign-inverted velocities $\boldsymbol{v}$. $\boldsymbol{r}$ are Cartesian coordinates. (Note that because of a printing error, the underlines are missing in Ref. 5.) For typical condensed-phase systems, the committors (averaged over velocities $\boldsymbol{v}$ in the case of deterministic dynamics) are close to either 0 or 1 nearly everywhere, except in a small transition region. Therefore, the products of committors, $\phi_A\phi_B$, are near zero almost everywhere.

To cope with the resulting problem of low acceptance, Dellago *et al.*[4] proposed to use phase-space dynamics and minimally perturb the momentum at the shooting point. However, for systems with long TPs, typical of many problems in condensed phase, memory of the initial momentum is rapidly lost. In forward-flux sampling[6] and transition interface sampling,[7] the problem is addressed by staging the sampling through a sequence of non-intersecting dividing surfaces that foliate the configuration space between the two metastable states. The "aimless shooting" algorithm by Peters and Trout[8] is an elegant approach that operates within the standard TP shooting

a)Author to whom correspondence should be addressed: gerhard.hummer@biophys.mpg.de

framework without using order parameters. Trial trajectories are shot off from the point where shooting was last successful or from points a certain number of frames before or after in time. Recently, Brotzakis and Bolhuis[9] built on the idea of biasing the choice of the starting point and extended it to one-way shooting for stochastically thermostatted dynamics. Earlier, Juraszek and Bolhuis[10] had combined transition interface sampling with a bias on the shooting points to improve the sampling efficiency for protein folding simulations. In an alternative approach, one of us proposed to shoot TPs from a single dividing surface,[5] with random directions, as in aimless shooting,[8] but requiring an order parameter as in transition interface sampling.[7] Ideally, the dividing surface is chosen to capture the ensemble of transition states at the top of the kinetically relevant barriers.[5] This approach promises to enhance the efficiency of sampling, but it requires a well chosen dividing surface, a good sampling of initial conditions on that surface, and an elaborate accounting for the relative weights of TPs.

In many practical cases, one has at least partial information about the mechanism in the form of reaction coordinates, as witnessed by the success of path sampling methods such as forward-flux sampling[6] and transition interface sampling,[7] and of the simpler Markovian form of milestoning.[11] For instance, despite the evident complexity in the folding of small proteins, as seen in simulations,[12] it was found that the fraction of native contacts forms an excellent reaction coordinate.[13,14] Our aim is to use such information to bias the choice of shooting points.

We extend the procedure of "shooting from the top" by initiating trial TPs recursively from a pre-defined "shooting range" of finite width instead of a dividing surface. In this way, we aim to meet the competing demands of high acceptance, efficient sampling over a broad range of TP families, and simple accounting for the weights in the TP ensemble. To avoid getting trapped in a single transition saddle (or tube), one can combine "shooting from the top" with regular TP shooting by using weighted shooting ranges. The resulting algorithm requires minimal bookkeeping and is applicable to a wide range of dynamics, including Newtonian dynamics with and without certain thermostats, Langevin dynamics in phase space, overdamped Langevin (or Brownian) dynamics, and Metropolis Monte Carlo (MC) sampling. This type of shooting is particularly useful where one has dynamically relevant order parameters,[15] i.e., good reaction coordinates, and where the cost of generating TPs is high (e.g., in protein folding or conformational dynamics).

The paper is organized as follows. We first derive the general algorithm of "TP shooting from the top." We then illustrate its application and characterize its efficiency for simple model systems. We also show how the shooting range can be adaptively optimized to guide it toward the transition state ensemble. To conclude, we discuss possible applications as well as limitations and ways to address them.

II. THEORY

A. Path probability

In the following, we assume that trajectories are created by integrating the equations of motions over discrete and uniform intervals of time δt. Determined by the frequency of saving intermediate structures, δt is usually a multiple of the integration time step Δt (or a certain number of MC steps). We further assume that in the full phase space, the dynamics is Markovian, not explicitly time dependent, microscopically time reversible, and preserving an equilibrium distribution. These conditions are satisfied by a wide range of dynamics used in computer simulations, including deterministic Newtonian dynamics on a static potential surface, Langevin dynamics in phase space, overdamped Langevin dynamics, and a variety of thermostatted dynamics. For dynamics with inertia, the multidimensional phase points $x = (\boldsymbol{r}, \boldsymbol{v})$ include velocities $\boldsymbol{v}$. For Brownian or MC dynamics, we consider only coordinates $x = \boldsymbol{r}$. Note that x is the state of the entire system and also includes solvent degrees of freedom (and, in the extended Lagrangian dynamics, the additional degrees of freedom).

For a given dynamics, the probability of going from one phase point to the next in time δt defines the propagator or Green's function $p(x_i \to x_{i+1}) \equiv p(x_{i+1}, \delta t | x_i, 0)$, where $\int dx_{i+1} p(x_i \to x_{i+1}) = 1$. The probability p of the time reversed dynamics then satisfies the detailed balance condition[4]

$$p_{\mathrm{eq}}(x_{i+1}) p(\underline{x}_{i+1} \to \underline{x}_i) = p_{\mathrm{eq}}(x_i) p(x_i \to x_{i+1}), \tag{1}$$

where $\underline{x} = (\boldsymbol{r}, -\boldsymbol{v})$ denotes phase points with sign-inverted velocities, and $p_{\mathrm{eq}}(x) = p_{\mathrm{eq}}(\underline{x})$ is the conserved equilibrium distribution.

Our aim is to sample trajectories from the TP ensemble. TPs connect two non-overlapping regions A and B in phase space directly. These regions are usually chosen to cover distinct metastable states with the boundaries of A and B defined in terms of one or several order parameters. The probability of a path $X = (x_0, x_1, \ldots, x_N)$ in the TP ensemble is then proportional to[3,5]

$$p[X] \propto p_{\mathrm{eq}}(x_0) \left(\prod_{i=0}^{N-1} p(x_i \to x_{i+1}) \right) \times \left(\prod_{j=1}^{N-1} h_C(x_j) \right) \left(h_A(x_0) h_B(x_N) + h_B(x_0) h_A(x_N) \right), \tag{2}$$

where we denote paths with capital letters and the arguments of path probabilities with square brackets to distinguish them from phase-space probability densities and transition probabilities. The indicator functions are defined as $h_M(x) = 1$ for all $x \in M$ and $h_M(x) = 0$ otherwise, for $M = A$, B, and C, with C as the phase-space region outside both A and B. Note that here we use the TP definition of Ref. 5 for paths of variable duration contained entirely in the transition region C (except for the first and last points, with the paths truncated before and after, respectively). That is, the last term accounts for the possibility of A-to-B and B-to-A paths.

B. Transition path shooting

In its most general form, we select a shooting point x_i from an existing path $X = (x_0, x_1, \ldots, x_N)$ according to a probability $p_{\mathrm{sel}}(x_i|X)$, map it to another phase-space point x' according to a probability density $p_{\mathrm{pert}}(x'|x_i)$, and then propagate

trajectories first forward and then backward in time from x' until either A or B are reached. If the forward and backward trajectory segments end in opposite regions, A and B or B and A, then the new path $X' = (x'_0, x'_1, \ldots, x'_{N'})$, with $x'_j = x'$ for some j by construction, forms a candidate TP. To time-order paths so that they consistently go from A to B, the segment ending in A is time reversed (i.e., re-ordered from end to beginning) and placed at the beginning of the TP. This trial path is then accepted or rejected according to a criterion that ensures paths are sampled according to Eq. (2), using the Metropolis Markov chain methodology (see Fig. 1). For this procedure to sample from the correct distribution, the usual requirements for Markov chain sampling have to be satisfied. In particular, the TP space has to be connected, i.e., every path should be reachable in principle from every other path in a finite number of steps. By repeatedly applying this procedure, a sample of representative TPs is generated.

In the following, we derive the acceptance criterion of candidate TPs. In this derivation, we follow earlier studies[1,4,16] with minor adjustments to account for the choice of a shooting range. In particular, we require that path ensemble probabilities, Markov-chain generation probabilities, and acceptance probabilities are connected by detailed balance,

$$p[X]p_{\mathrm{gen}}[X \to X']p_{\mathrm{acc}}[X \to X'] = p[X']p_{\mathrm{gen}}[X' \to X]p_{\mathrm{acc}}[X' \to X]. \quad (3)$$

Since trajectory segments are generated by time integration, the generation probability can be written in terms of the propagator

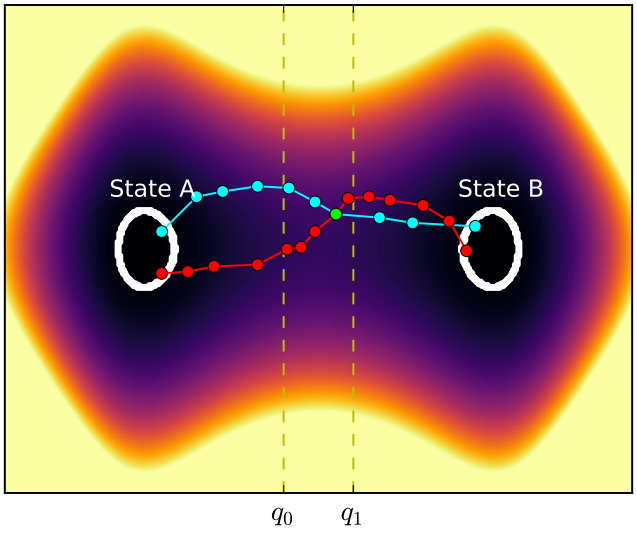

FIG. 1. Schematic illustration of transition path "shooting from the top." The last accepted TP trajectory X (cyan) and the newly generated trial trajectory X' (red) are shown on top of a two-dimensional free energy landscape (where dark colors indicate low free energies). The shooting point (green) lies on both trajectories. The metastable regions A and B are indicated (white outlines) as well as the "shooting range" within boundaries $q(x) = q_0$ and $q(x) = q_1$ (dashed yellow lines). Trajectories X and X' have n = 3 and $n' = 5$ points within the shooting range. The Metropolis acceptance probability to go from X to X' is thus 3/5 = 60%. Conversely, the acceptance probability to go from X' to X is 100% because $n < n'$.

$$p_{\mathrm{gen}}[X \to X'] = p_{\mathrm{sel}}(x_i|X)p_{\mathrm{pert}}(x'_j|x_i) \times \prod_{k=j}^{N'-1} p(x'_k \to x'_{k+1}) \prod_{m=0}^{j-1} p(\underline{x}'_{m+1} \to \underline{x}'_m). \quad (4)$$

Here and in the following, we dropped the indicator functions $h_M(x_k)$ since in practice those are accounted for in an initial rejection of all trial paths that are not candidate TPs. Using Eq. (1) for $p[X]$, we arrive at

$$p_{\mathrm{gen}}[X \to X'] = \frac{p_{\mathrm{sel}}(x_i|X)p_{\mathrm{pert}}(x'_j|x_i)}{p_{\mathrm{eq}}(x'_j)} p[X']. \quad (5)$$

After substituting this and an analogous expression for $p_{\mathrm{gen}}[X' \to X]$ into Eq. (3), we obtain an expression for the ratio of acceptance probabilities,

$$\frac{p_{\mathrm{acc}}[X \to X']}{p_{\mathrm{acc}}[X' \to X]} = \frac{p_{\mathrm{sel}}(x'_j|X')p_{\mathrm{pert}}(x_i|x'_j)p_{\mathrm{eq}}(x'_j)}{p_{\mathrm{sel}}(x_i|X)p_{\mathrm{pert}}(x'_j|x_i)p_{\mathrm{eq}}(x_i)}, \quad (6)$$

in terms of ratios of the probabilities $p_{\mathrm{sel}}(x'_j|X')$ and $p_{\mathrm{sel}}(x_i|X)$ of selecting candidate shooting points from existing paths X and X', of the probability densities $p_{\mathrm{pert}}(x_i|x'_j)$ and $p_{\mathrm{pert}}(x'_j|x_i)$ of creating perturbed starting points from the candidate shooting points, and of their equilibrium probability densities $p_{\mathrm{eq}}(x'_j)$ and $p_{\mathrm{eq}}(x_i)$. The Metropolis acceptance criterion

$$p_{\mathrm{acc}}[X \to X'] = \min\left[1, \frac{p_{\mathrm{sel}}(x'_j|X')p_{\mathrm{pert}}(x_i|x'_j)p_{\mathrm{eq}}(x'_j)}{p_{\mathrm{sel}}(x_i|X)p_{\mathrm{pert}}(x'_j|x_i)p_{\mathrm{eq}}(x_i)}\right] \quad (7)$$

satisfies this condition. Note that, for notational simplicity, in Eqs. (6) and (7) we did not include the products of indicator functions imposing that paths X and X' are transition paths by having one of their two endpoints in A, the other in B, and all other points in C, i.e., neither in A nor in B. As in regular Metropolis Monte Carlo sampling, after rejecting a shooting attempt that produced no TP or a TP of low weight, the last accepted path is added once more to the growing TP ensemble.

In the simplest form of sampling by "shooting from the top," we define a shooting range S within which we choose candidate shooting points. A similar formulation has been used in S-shooting, a reactive-flux type rate calculation introduced by Menzl *et al.*[17] In the following, we assume that S is defined as $S = \{x|q_0 < q(x) < q_1\}$ in terms of a single scalar order parameter $q(x)$. However, more complex definitions are possible, including a combination of weighted shooting ranges, as described below. Usually, the region bounded by q_0 and q_1 will not overlap with the metastable regions A and B, but this is not a requirement. Importantly, though, it is required that all TPs pass through S at least once. For a given path X, let the set $X_{\mathrm{sr}} = \{x_k|x_k \in X;\ q_0 < q(x_k) < q_1\}$ contain all of its $n = |X_{\mathrm{sr}}|$ points in the shooting range. Among these points, we choose one at random, i.e., $p_{\mathrm{sel}}(x_i|X) = 1/n$ if $x_i \in X_{\mathrm{sr}}$ and 0 otherwise. On the candidate TP X', the corresponding set $X'_{\mathrm{sr}} = \{x'_k|x'_k \in X';\ q_0 < q(x'_k) < q_1\}$ contains $n' = |X'_{\mathrm{sr}}|$ points, and $p_{\mathrm{sel}}(x'_j|X') = 1/n'$ if $x'_j \in X'_{\mathrm{sr}}$ and 0 otherwise. For simplicity, we generate perturbed shooting points x'_j from x_i and x_i from x'_j such that $p_{\mathrm{pert}}(x_i|x'_j)p_{\mathrm{eq}}(x'_j) = p_{\mathrm{pert}}(x'_j|x_i)p_{\mathrm{eq}}(x_i)$. For stochastic dynamics, such as Langevin dynamics, we simply set $x'_j = x_i$, without perturbation, and rely on the

newly generated sequence of random numbers to create divergent trajectories. For deterministic dynamics, we modify the velocities in a way that is reversible, e.g., by random rotation or by redrawing Maxwell-Boltzmann velocities that are then rescaled to preserve the kinetic energy exactly. For such perturbations, the acceptance probability Eq. (7) for the new path simplifies to

$$p_{\rm acc}[X \to X'] = \min\left[1, \frac{n}{n'}\right], \tag{8}$$

i.e., a candidate path is accepted with probability one if it has fewer points within the shooting range and with a probability equal to the ratio n/n' of the number of points in the shooting range if $n \le n'$. In this way, we correct for the bias towards paths that dwell for long times in the shooting range. This procedure can be thought of as a generalization of the procedure of sampling transition paths of unequal length,[5,7] where $n = N$ and $n' = N'$. As a further generalization, one can select shooting points according to a weighting function, which will be discussed below. In this context, we also refer to a recent review by Bolhuis and Dellago,[16] in which a biased choice of shooting points has been discussed and the corresponding acceptance criterion has been derived.

A minor technical issue is that for very narrow shooting ranges, newly created paths saved at long intervals δt have only one discrete point x_i (the shooting point) in the range even if they cross it multiple times. One can then widen the shooting range, increase the saving frequency, or use a wider $p_{\rm pert}(x'|x)$, in which also the Cartesian coordinates $\boldsymbol{r}$ are perturbed. A more challenging problem is that TP space may be nearly disjoint, e.g., in cases where a reaction can proceed with two distinct mechanisms. In the above algorithm, candidate shooting points are ideally concentrated in the saddle regions, which could make it difficult to traverse from one saddle to another through repeated accepted shooting moves. In practice, this may be less of a problem since the choice of the order parameter $q(x)$ will hardly ever be ideal. Nonetheless, as a counter measure, one can modify $p_{\rm sel}(x|X')$, e.g., by shooting with probability w from the shooting range S and with probability $1 - w$ from the entire path (such that $n = N$ and $n' = N'$ in the latter case). Below, we describe a generalization of this idea to weighted shooting ranges.

C. Sampling efficiency

Using TP theory,[5,18] we can determine the expected probabilities of generating a TP in a shooting move. The points on TPs have a probability density in phase space given by[5]

$$p(x|{\rm TP}) = \frac{p_{\rm eq}(x)[\phi_A(x)\phi_B(\underline{x}) + \phi_B(x)\phi_A(\underline{x})]}{\int dx' p_{\rm eq}(x')[\phi_A(x')\phi_B(\underline{x}') + \phi_B(x')\phi_A(\underline{x}')]}, \tag{9}$$

where $\phi_B(x) = 1 - \phi_A(x)$. For a point x, the probability that shooting will lead to a TP (i.e., of hitting A and B) is $p({\rm hit}|x) \equiv p({\rm TP}|x) = \phi_A(x)\phi_B(\underline{x}) + \phi_B(x)\phi_A(\underline{x})$. The probability $p({\rm hit})$ of creating a TP in "shooting from the top," averaged over points x on existing TPs and in the shooting range, is thus

$$p({\rm hit}) = \frac{\left\langle h_S(x)[\phi_A(x)\phi_B(\underline{x}) + \phi_B(x)\phi_A(\underline{x})]^2\right\rangle_{\rm eq}}{\left\langle h_S(x)[\phi_A(x)\phi_B(\underline{x}) + \phi_B(x)\phi_A(\underline{x})]\right\rangle_{\rm eq}}, \tag{10}$$

with $\langle \ldots \rangle_{\rm eq}$ being an average over the equilibrium distribution. $h_S(x)$ is the indicator function for the shooting range, i.e., $h_S(x) = 1$ if $q_0 < q(x) < q_1$, and $h_S(x) = 0$ otherwise. [Note that for simplicity, we did not include $p_{\rm pert}(x'|x)$, which would require an additional average.]

For dynamics that is Markovian in configuration space with coordinates $\boldsymbol{r}$, these expressions simplify to

$$p(\boldsymbol{r}|{\rm TP}) = \frac{p_{\rm eq}(\boldsymbol{r})\phi_A(\boldsymbol{r})\phi_B(\boldsymbol{r})}{\int d\boldsymbol{r}' p_{\rm eq}(\boldsymbol{r}')\phi_A(\boldsymbol{r}')\phi_B(\boldsymbol{r}')} \tag{11}$$

and

$$p({\rm hit}) = \frac{2\left\langle h_S(\boldsymbol{r})[\phi_A(\boldsymbol{r})\phi_B(\boldsymbol{r})]^2\right\rangle_{\rm eq}}{\langle h_S(\boldsymbol{r})\phi_A(\boldsymbol{r})\phi_B(\boldsymbol{r})\rangle_{\rm eq}}. \tag{12}$$

For the inertial dynamics without recrossings assumed in transition state theory, and for x a transition state, the terms in square brackets of Eq. (10) are equal to one. The maximum acceptance probability is thus $p({\rm hit}) = 1$ if $h_S(x)$ captures the transition states exactly and exclusively. By contrast, for diffusive dynamics, the acceptance probability reaches its maximum at $p({\rm hit}) = 1/2$, again if $h_S(x) = 1$ only for transition states with $\phi_A(x) = \phi_B(x) = 1/2$. If the shooting range extends beyond the region of transition states, $p({\rm hit})$ is reduced. If one assumed that all TPs are accepted and the correlation between consecutively sampled TPs is independent of the extent of the shooting range (which is unlikely to be the case in practice), the relative efficiency of sampling with different shooting ranges would be proportional to $p({\rm hit})$ in Eq. (10) or (12).

Importantly, even if the shooting range is defined suboptimally, the acceptance probability remains reasonably high. The reason is that shooting is initiated from points distributed according to $p(x|{\rm TP})$ and not $p_{\rm eq}(x)$. As shown in Eq. (9), the TP ensemble is already strongly biased towards transition states. This bias helps explain why standard TPS with shooting moves[4] works reasonably well even without restricting the shooting range.

D. Weighted shooting ranges and adaptive optimization

So far, we have assumed a single shooting range within which candidate shooting points are chosen at random. In the following, we extend this approach to using a combination of weighted shooting ranges. We assume that we have a suitably chosen weight function $w(x)$ according to which the points should be chosen with $w(x) \ge 0$. The probabilities to select $x_i \in X$ and $x'_j \in X'$ are then

$$p_{\rm sel}(x_i|X) = \frac{w(x_i)}{\sum_{x_k \in X} w(x_k)}, \tag{13}$$

$$p_{\rm sel}(x'_j|X') = \frac{w(x'_j)}{\sum_{x'_k \in X'} w(x'_k)}. \tag{14}$$

These expressions can be substituted into Eqs. (6) and (7) to obtain the corresponding acceptance criteria. In this way, one can, e.g., predominantly select candidate points from a core shooting range [where $w(x)$ is large] and occasionally from the entire region outside A and B [where $w(x)$ is small]. Biased

choices of shooting points as a means to improve the sampling efficiency have been discussed before by Bolhuis and Dellago.[16]

The weight function $w(x)$ can be adaptively optimized during TPS. In practice, it will usually be formulated as a function of one or several order parameters Q_1, Q_2, . . ., i.e., $w = w[Q_1(x), Q_2(x), \ldots]$. In Sec. V, we sketch possible machine learning and reaction-coordinate based optimizations. Here, we pursue an even simpler strategy in which the shooting range is moved and its width is adjusted. In effect, this algorithm maximizes the overlap of the shooting range with the ensemble of points where shooting has been particularly successful. Since the latter can be broadly identified as "transition states," the algorithm is related to procedures to optimize reaction coordinates. However, such procedures are usually applied *after* simulations are completed on the basis of transition path and equilibrium ensembles.[5,8,15,19] The optimization strategy used here is also related to the aimless shooting algorithm by Peters and Trout,[8] which aims to concentrate shooting attempts near the dividing surface *without* invoking an order parameter. As a possible trade-off, by choosing the shooting point at or near the last successful shooting point, aimless shooting produces sequences of TPs that tend to share a common point or pass through a narrow region. By contrast, if a sufficiently wide shooting range is used, shooting points can be widely separated, which should accelerate the decorrelation of TPs and facilitate the crossing between different transition states.

We define the shooting range in terms of an order parameter $q(\boldsymbol{r})$. The aim is to optimize this definition on the basis of the collected statistics for $p(\text{hit}|\boldsymbol{r})$, i.e., the probability, that a TP shooting attempt initiated at $\boldsymbol{r}$ is successful in the sense of producing a TP. We optimize both the location and the size of the shooting range. In a shifting move, we move the center of the shooting range towards the weighted center of $p(\text{hit}|\boldsymbol{r})$ calculated over the current shooting range. In a narrowing move, we check if it would be beneficial to narrow down (or widen) the shooting range symmetrically from both sides. Specifically, we test if

$$\frac{\int_{\text{trialSR}} d\boldsymbol{r}\, p(\text{hit}|\boldsymbol{r})}{\int_{\text{SR}} d\boldsymbol{r}\, p(\text{hit}|\boldsymbol{r})} \leqslant \frac{A_{\text{trialSR}}}{A_{\text{SR}}},$$

where A_{trialSR} and A_{SR} stand for measures of the areas of the trial and existing shooting ranges. In our one dimensional case, A is proportional to the width of the shooting range. If one has a set of order parameters that in (linear) combination define a decent reaction coordinate, straightforward extensions of the iterative shooting range optimization procedure outlined above can be used to translate, rotate, and narrow or widen the shooting range to capture as much of the transition state ensemble as possible. More general approaches, which change the order parameter itself, are discussed in Sec. V.

III. MODEL SYSTEMS

We illustrate and test the TP shooting algorithm in applications to simple model systems. The validity of the ensemble of reactive trajectories generated with the algorithm is tested by comparison with the ensemble of reactive trajectories harvested from long equilibrium trajectories.

As a numerical integrator for the overdamped Langevin equation, we used the Euler-Maruyama discretization scheme as implemented, e.g., in GROMACS,[20]

$$\boldsymbol{r}(t + \Delta t) = \boldsymbol{r}(t) + D\,\Delta t\,\boldsymbol{F}(\boldsymbol{r}(t))/k_BT + \sqrt{2D\Delta t}\boldsymbol{g},$$

where $\boldsymbol{r}(t)$ and $\boldsymbol{r}(t + \Delta t)$ are the current and new positions, respectively, Δt is the time step, $\boldsymbol{F}(\boldsymbol{r})$ is the force evaluated at $\boldsymbol{r}$, D is the diffusion coefficient, k_B is Boltzmann's constant, T is the temperature, and $\boldsymbol{g}$ is a vector of uncorrelated Gaussian random numbers with zero mean and unit variance. This integrator is used for the generation of the equilibrium trajectories as well as in the corresponding TP sampling. Note that this simple integrator conserves the Boltzmann equilibrium distribution only approximately in a time step dependent manner.

To study inertial dynamics in phase space, we implemented a leap-frog type "impulsive Langevin" integrator,[21] which is also implemented in GROMACS.[20] Trajectories were advanced according to

$$\begin{aligned}
\boldsymbol{v}'\left(t + \frac{\Delta t}{2}\right) &= \boldsymbol{v}\left(t - \frac{\Delta t}{2}\right) + \frac{\boldsymbol{F}(\boldsymbol{r}(t))\Delta t}{m},\\
\Delta\boldsymbol{v} &= -\alpha\boldsymbol{v}'\left(t + \frac{\Delta t}{2}\right) + \sqrt{\frac{k_BT}{m}\alpha\,(2-\alpha)}\boldsymbol{g},\\
\boldsymbol{r}(t + \Delta t) &= \boldsymbol{r}(t) + \left(\boldsymbol{v}'\left(t + \frac{\Delta t}{2}\right) + \frac{\Delta\boldsymbol{v}}{2}\right)\Delta t,\\
\boldsymbol{v}\left(t + \frac{\Delta t}{2}\right) &= \boldsymbol{v}'\left(t + \frac{\Delta t}{2}\right) + \Delta\boldsymbol{v},\\
\alpha &= 1 - e^{-\gamma\Delta t},
\end{aligned}$$

where $\boldsymbol{v}$ is the velocity, γ is the friction rate (with dimension of inverse time), and $\boldsymbol{g}$ is again a vector of uncorrelated Gaussian random numbers, drawn at each time step, with zero mean and unit variance. In the following, we will refer to this time integration simply as "Langevin."

As a model system, we considered a two-dimensional potential surface

$$V(x, y) = B\left(\left(x^2 - 1\right)^2 + (x - y)^2\right),$$

where B is the barrier height in units of k_BT. (In the following, we will use reduced units, with mass $m = 1$, and energies in units of k_BT.) The two stable states are defined as all points (x, y) with $V(x, y) < 0.1B$ (see Fig. 2). The state referred to as A is located around $\boldsymbol{r} = (-1, -1)$, whereas the state B is located around $\boldsymbol{r} = (1, 1)$.

All probabilities were evaluated on a 2D grid with a spacing of 0.01 from −2 to +2 on both axes. The probability $p(\boldsymbol{r}|\text{TP})$ was represented as a normalized histogram of points along transition paths spaced uniformly in time.

Starting from an existing transition path, the shooting procedure for the overdamped case consists of choosing at random a point within the shooting range and then propagating two trajectory segments from there. If the two newly grown trajectory segments reach different stable states, A and B or B and A, the newly generated transition path is accepted according to

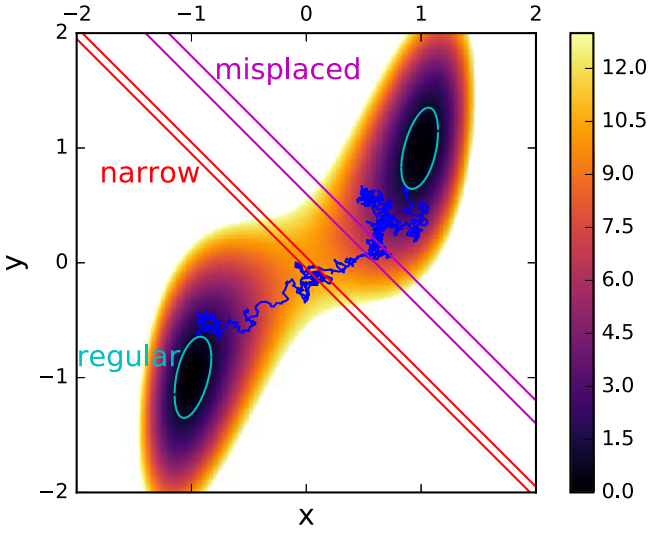

FIG. 2. Top view of the potential $V(x, y)$. The borders of the employed shooting ranges are shown in red (narrow), magenta (misplaced), and cyan (regular). Regular shooting uses the entire region outside the states A and B indicated with cyan boundaries. For illustration, a transition path generated with overdamped Langevin dynamics is shown in blue.

Eq. (7). The shooting procedure for the Langevin case consists of selecting a random point on the old transition path lying in the shooting range and propagating two trajectory segments in opposite directions in time by inverting the velocities of the "backwards" segment. Whereas the coordinates $\boldsymbol{r}$ of the shooting point are not changed, a new velocity vector $\boldsymbol{v}$ is drawn at random and rescaled to match the absolute value of the old velocity at the chosen shooting point. Therefore, the total kinetic energy does not change and the acceptance criterion stays simple.

Different choices of shooting ranges were employed to quantify the effects of various options (see Fig. 2). The shooting range referred to as "narrow" is centered at the dividing surface and is only 0.1 units wide. It is bordered by the two functions $x_0(y) = -y - 0.05$ and $x_1(y) = -y + 0.05$. The shooting range referred to as "misplaced" does not include the dividing surface but completely separates the two stable states. Its bordering functions are $x_0(y) = -y + 0.4$ and $x_1(y) = -y + 0.7$. The shooting range "regular" consists of all points outside of A and B. Using the "regular" range is thus equivalent to the traditional TP shooting algorithm without a tightened shooting range.[4]

We measure the efficiency of TP sampling with a particular shooting range as the ratio of generated transitions to the total number of MC steps. To quantify the difference between probability distributions $p_1(x, y)$ and $p_2(x, y)$, we use the non-symmetrized Kullback-Leibler (KL) divergence

$$D_{\mathrm{KL}}(p_1||p_2) \equiv \sum_{i,j} p_1(x_i, y_j) \ln \frac{p_1(x_i, y_j)}{p_2(x_i, y_j)}, \tag{15}$$

where the sum is over the bins i and j and the histograms are normalized, $\sum_{i,j} p_1(x_i, y_j) = \sum_{i,j} p_2(x_i, y_j) = 1$. To compare the TP densities, $p(\boldsymbol{r}|\mathrm{TP})$, we use $D_{\mathrm{KL}}(\mathrm{EQ}||\mathrm{SR})$, i.e., the well-sampled TP density from shooting with a shooting range (SR) is used as reference p_2 and the TP densities from long equilibrium (EQ) runs are used as p_1. We also compare the distribution of transition path times, i.e., the time on a TP between the last point within A and the first point within B.

IV. RESULTS AND DISCUSSION

A. Transition path ensembles

TP sampling reproduced the equilibrium ensemble of transition paths, as harvested from long equilibrium trajectories, for all shooting ranges, barrier heights, and integrators. This can be best seen for a moderately high barrier ($B = 3\,k_BT$) because there the equilibrium trajectory easily contains enough transitions for a meaningful statistics. As shown in Fig. 3(c), there are no systematic deviations between the distributions $p(\boldsymbol{r}|\mathrm{TP})$ from TP sampling and from the equilibrium trajectory. The KL divergences range from $D_{\mathrm{KL}}(\mathrm{EQ}||\mathrm{narrowSR}) = 0.0019$ to $D_{\mathrm{KL}}(\mathrm{EQ}||\mathrm{regular}) = 0.0038$ and show excellent agreement between the sampled and equilibrium distributions (see Table I). The distribution of transition times, $p(t_{\mathrm{TP}})$, was also correctly reproduced for all choices of shooting ranges [see Fig. 3(d) for $B = 3\,k_BT$ and Table I for other systems].

Similar results were obtained with the impulsive Langevin integrator without noticeable systematic derivations between the sampled and equilibrium distributions (Fig. S1 of the supplementary material). Here the Kullback-Leibler divergences are a bit higher (Table I) but still indicate that the distributions are in excellent agreement.

For high barriers ($B = 10\ k_BT$), where the TP sampling becomes increasingly interesting because of the growing disparity of time scales between waiting time and transition time, the distributions $p(\boldsymbol{r}|\mathrm{TP})$ and $p(t_{\mathrm{TP}})$ were also correctly reproduced but a comparison with the equilibrium trajectory became difficult because it only contained a few transitions. The Kullback-Leibler divergences are therefore a bit higher than for the lower barrier but still indicate excellent agreement between the distributions with values from $D_{\mathrm{KL}}(\mathrm{EQ}||\mathrm{narrowSR}) = 0.039$ to $D_{\mathrm{KL}}(\mathrm{EQ}||\mathrm{misplacedSR}) = 0.061$.

As can be seen from Fig. S2(C) of the supplementary material, the different TP sampling schemes produce distributions that deviate by about the same amount from the equilibrium distribution. These difference values are calculated at every TP shooting attempt as the sum of the absolute values of the difference between the distributions $p(\boldsymbol{r}|\mathrm{TP})$ from TP sampling at that particular step and the corresponding distribution from the complete equilibrium trajectory. Therefore, it seems likely that this difference is due to the fact that the number of equilibrium transitions is too small to arrive at a good equilibrium distribution. This is supported by the agreement of the distributions of transition times obtained from the equilibrium trajectory and from TP sampling with different shooting ranges [Fig. S2(D)].

B. Efficiency

The increase in efficiency in generating new transitions through the definition of a shooting range varies naturally depending on the shooting range location. The shooting range "narrow," which is centered at the dividing surface, reached almost the ideal efficiency of 0.5 for the overdamped case independent of the barrier height. By contrast, for high barriers, the "misplaced" shooting range yielded an efficiency lower than

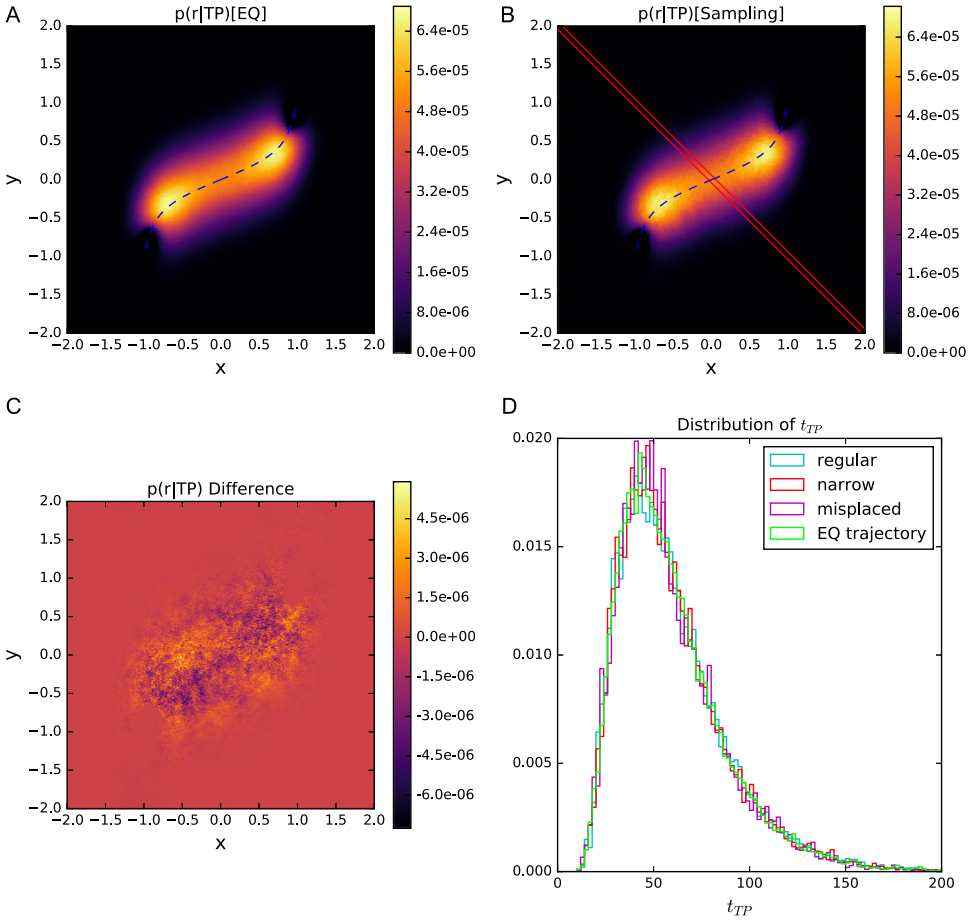

FIG. 3. Comparison of transition paths from TP shooting and equilibrium simulations. Results are shown for the overdamped Langevin integrator with a barrier $B = 3\ k_BT$, a time step $\Delta t = 0.01$, and a diffusion coefficient $D = 0.01$. Points were saved for analysis at every time step (i.e., $\Delta t = \delta t$). (Top) Normalized probability densities $p(\boldsymbol{r}|\mathrm{TP})$ from (a) equilibrium trajectories and (b) TP shooting. The path of steepest descent is marked by a blue dashed line. The shooting range is marked by red lines in panel (b). (c) Difference in $p(\boldsymbol{r}|\mathrm{TP})$ from equilibrium and TP shooting. (d) Normalized histograms of the transition times t_{TP}.

the "normal" shooting range. The reason is that for high barriers, a shot initiated far from the dividing surface has almost no chance of generating a transition, whereas with a wide shooting range, the occasional "lucky draw" of a point near the dividing surface dominates the efficiency in generating transitions.

While the efficiency of the narrow shooting range stays almost constant as the barrier height increases (0.48 for $B = 3\ k_BT$ and 0.46 for $B = 10\ k_BT$), the efficiency with the "regular" shooting range drops with increasing barrier height from 0.24 ($B = 3\ k_BT$) to 0.11 ($B = 10\ k_BT$). This makes it clear that the definition of a shooting range becomes increasingly important with increasing barrier heights.

It should also be noted that the efficiency in generating new transitions with the Langevin dynamics is generally a bit higher and increases with decreasing friction. The reason is the partial conservation of momentum: a pair of trajectories initiated with opposite velocities at the barrier top initially fall off the barrier in opposite directions and are likely to continue their path toward opposite states.

C. Iterative shooting range optimization

A misplaced shooting range can result in low efficiencies (see Fig. 4). To address this problem, we devised an algorithm that is able to iteratively optimize the location of the shooting range. After a fixed number of shots, we first tried a narrowing move and, if no narrowing took place, a shifting move. To retain the correct weighting between the

TABLE I. Kullback-Leibler divergences between the $p(\boldsymbol{r}|\mathrm{TP})$ distributions sampled by TP shooting for a given shooting range and by equilibrium simulations. The number of unique transition paths in the respective TP ensemble is listed in parentheses. For comparison, the same number of shooting attempts was made in every setup independent of the shooting range.

System	Equilibrium	Narrow SR	Misplaced SR	Regular
Overdamped Langevin, $B = 3\ k_BT$	(28 941)	0.001 85 (24 990)	0.003 61 (12 688)	0.003 82 (17 770)
Langevin, $B = 3\ k_BT$, $\gamma = 50$	(7 268)	0.007 22 (12 606)	0.011 33 (6 211)	0.008 21 (8 951)
Langevin, $B = 3\ k_BT$, $\gamma = 20$	(17 380)	0.008 71 (12 422)	0.015 01 (6 263)	0.009 51 (8 503)
Overdamped Langevin, $B = 10\ k_BT$	(605)	0.038 6 (50 326)	0.061 2 (2 238)	0.043 2 (17 986)
Langevin, $B = 10\ k_BT$, $\gamma = 50$	(430)	0.047 7 (24 662)	0.095 7 (1 293)	0.051 7 (8 828)
Langevin, $B = 10\ k_BT$, $\gamma = 20$	(492)	0.079 7 (25 475)	0.125 3 (1 166)	0.086 2 (9 148)

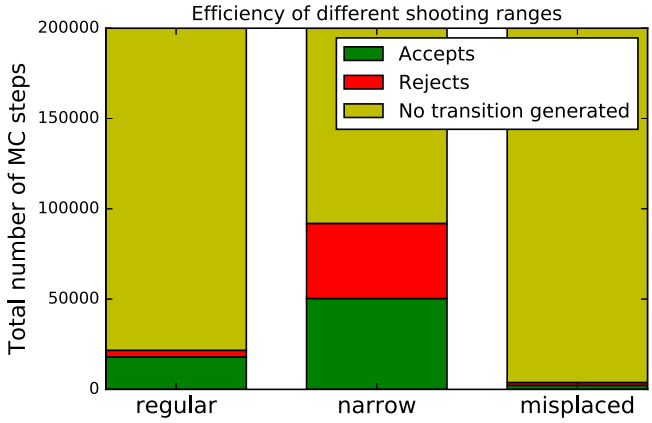

FIG. 4. Decomposition of the 200 000 MC shooting attempts into accepts, rejects, and no transition generated depending on the choice of the shooting range. The parameters used are $B = 10\ k_BT$, $\Delta t = 0.03$, and $D = 0.01$ as for Fig. S2 of the supplementary material.

transitions (and thereby generate an equilibrium ensemble of transitions), the number of points in the shooting range has to be recalculated for the current transition path after the shooting range is changed. If none of the saved points on the current TP is in the trial shooting range, the shooting range is not changed.

We have tested different numbers of shots between the optimizing steps and have found that between 25 and 50 shots are sufficient to improve the location of the shooting range significantly. The algorithm then reached an optimal shooting range after 15-30 optimizing steps. While in general less optimizing steps are needed if more shots are undertaken between two successive optimizations, it is beneficial to move the shooting range as frequently as possible to make use of the increased efficiency in generating TPs. In production TP sampling, an initially optimized shooting range should be kept fixed to create a properly weighted TP ensemble.

Figure 5 shows the result of a shooting range optimization every 50 shots, starting from a completely misplaced shooting range $x + y \in [-1.5, -0.5]$. The algorithm reached an optimal shooting range $[-0.02, 0.02]$ after 16 optimizing steps (800 shots). When changing the optimization frequency to every 25 shots, the algorithm reached the optimal shooting range after 34 optimizing steps (850 shots).

V. CONCLUDING REMARKS

We described an algorithm that uses shooting ranges to improve the efficiency in transition path sampling of rare events. The key idea is that shooting off trajectories from points in the transition state region is more likely to produce transition paths than shooting trajectories from points close to the reactant and product states. Similar reasoning has been used before to shoot trajectories from a well chosen dividing surface[5] or near points where previous shots were successful.[8,9] In a more general form, one can choose shooting points according to a weight function $w(x)$ in configuration (or phase) space, which would typically be defined in terms of an order parameter and adopt large values in the transition state region.

The weight function could be adaptively optimized to improve the sampling efficiency, e.g., by machine learning techniques applied to the shooting statistics. Neural networks have proven to be particularly flexible and efficient in representing functions of many variables, including many-body potential surfaces[22] and committor functions.[23] The latter approach, possibly extended to build on recent advances in machine learning,[24] could prove useful here. In essence, for given configurations $\boldsymbol{r}$, we want to obtain an accurate estimate of the committor $\phi_B(\boldsymbol{r}) \approx \varphi_B[Q_1(\boldsymbol{r}), Q_2(\boldsymbol{r}), \ldots]$ expressed as a function of scalar order parameters Q_i (e.g., distances, angles, and coordination numbers). The idea is to learn the function φ_B from the statistics of shots collected during the preceding TP sampling. The approximation φ_B to the committor, updated occasionally, can then be used to select shooting points with a bias toward points with high probability $p(\mathrm{TP}|\boldsymbol{r}) \approx 2(1-\varphi_B)\varphi_B$ of creating transition paths.

An adaptive version of the reaction coordinate optimization by Peters and Trout[8] provides a simple and powerful alternative to supervised learning approaches. These authors assume that for a given scalar reaction coordinate $Q = Q(\boldsymbol{r}|\boldsymbol{\alpha})$ parametrized in terms of a vector of free parameters $\boldsymbol{\alpha}$, one can write the committor as a sigmoidal function, $\varphi_B(Q) = [1 + \tanh(Q)]/2$. Peters and Trout[8] then maximized the likelihood

$$L = \prod_{\boldsymbol{r}_i \to B} \varphi_B\left[Q(\boldsymbol{r}_i)\right] \prod_{\boldsymbol{r}_j \to A} \left(1 - \varphi_B[Q(\boldsymbol{r}_j)]\right) \tag{16}$$

with respect to the free parameters $\boldsymbol{\alpha}$ in $Q = Q(\boldsymbol{r}|\boldsymbol{\alpha})$, where the products are over individual shots ending in B and A,

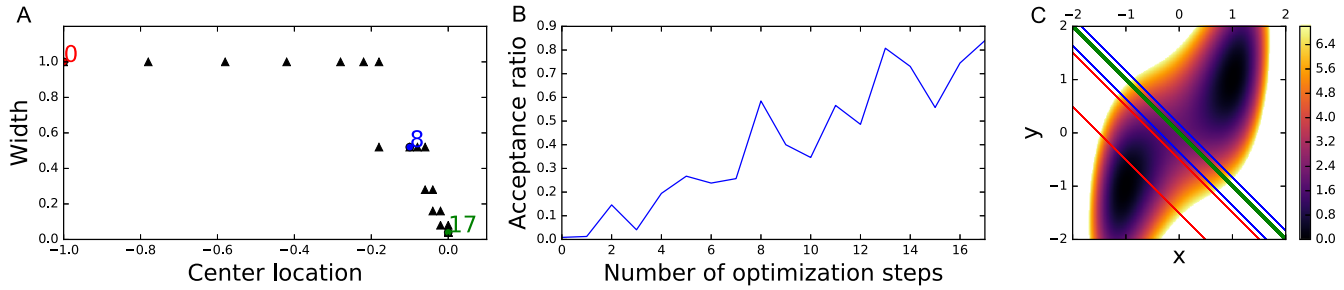

FIG. 5. Iterative optimization of the shooting range. Results are shown for Langevin dynamics with $B = 2\ k_BT$ and $q = x + y$ as a collective variable. Optimization attempts are made every 50 shots. (a) Center and width of the shooting range. Attempts 0, 8, and 17 are labeled. (b) Probability of generating and accepting transition paths in random shots, as the optimization progresses from step 0 to 17. (c) Shooting range boundaries (lines) on top of the 2D energy surface for the initial (step 0, red), an intermediate (step 8, blue), and the final optimal shooting range (step 17, green).

respectively. Building on this idea and making the optimization procedure adaptive, one would reoptimize the parameters $\boldsymbol{\alpha}$ at regular intervals. Incorporating the outcome of all preceding shooting attempts in the likelihood maximization should improve the quality both of Q as a reaction coordinate and of $\varphi_B(Q)$ as an estimator of the committor. In turn, $\varphi_B(Q)$ can then be used to bias the choice of shooting points, e.g., by setting the weight function in Eqs. (13) and (14) to $w(\boldsymbol{r}) \propto [\varphi_B[Q(\boldsymbol{r})](1 - \varphi_B[Q(\boldsymbol{r})])]^\nu$ with a power $\nu > 0$. In this way, one would adaptively optimize the reaction coordinate Q and use the improved coordinate to enhance the sampling of new transition paths. Such iterative optimizations amount to an attempt of localizing the transition state ensemble and are thus closely related to reaction coordinate optimization methods.[8,15,19] The difference here is that the relations[5,8,18] between the success in shooting and the committor and between the committor and $p(\mathrm{TP}|\boldsymbol{r})$ are exploited directly and "on the fly." However, as in the "likelihood maximization" method,[8] care should be exercised when using Q as a reaction coordinate if the TP ensemble, over which Q is optimized in this procedure, differs substantially from the equilibrium ensemble. Estimating the true committor $\phi_B(\boldsymbol{r})$ as $\varphi_B[Q(\boldsymbol{r})]$ at a point $\boldsymbol{r}$ typical of the equilibrium ensemble, but not the TP ensemble, effectively requires an extrapolation with the usual caveats.

When shooting points are chosen with a strong bias, such as the restriction to a tight shooting range, a possible concern is that the sampling space becomes effectively discontinuous. As a result, one would get trapped in local minima in transition path space (which is in addition to the general concern in TP sampling of having missed local intermediate states[1]). To sample all possible and relevant mechanisms in an efficient manner, it is important not only to have a high rate of success in TP shooting but also to transition at least occasionally between different "reaction mechanisms" corresponding to separated saddles on the (free) energy surface. To avoid getting trapped, one can use a weighted combination of shooting ranges and select, at least occasionally, points away from the core shooting range. A simple solution in practice is to select, with a certain frequency, the shooting point at random in the entire transition region, i.e., as in normal TPS. Such shots facilitate crossing between distinct transition states. This particular protocol, and more complex ones with different shooting ranges, can be expressed in terms of the weight function $w(x)$ in Eqs. (13) and (14). The acceptance criterion, Eq. (7), requires no modification, if the correct probabilities, Eqs. (13) and (14), to select shooting points are used. Similarly, it is straightforward to adapt the acceptance criterion to one-sided shooting[9] for dynamics with a stochastic thermostat.

It may also be interesting to put "shooting from the top" into an historical context. The concept of shooting from a dividing surface, which is ideally chosen to cut through the barrier tops, features prominently in the reactive-flux based methods of estimating reaction rates developed by Bennett[25] and Chandler.[26] Daru and Stirling[27] extended this approach from a dividing surface to a "saddle domain" in which trajectories are initiated. More recently, Menzl, Singraber, and Dellago[17] introduced S-shooting, in which trajectories are initiated from a shooting range "S" as a generalization of the Bennett-Chandler approach. Their algorithm samples from the ensemble of all trajectories passing through S, whereas here we sample only transition paths. Possibly one of the earliest implementations of path sampling by shooting as a means to probe rare events was in the context of diffusion-influenced bimolecular reactions and protein-protein binding in particular.[28–30] Northrup, Allison, and McCammon[28] generated association pathways by shooting Brownian dynamics trajectories of one of the binding partners subject to interactions with the other, where the mobile partner was initiated at a given separation from the fixed one with random orientations. Except for the (small) bias on the orientation at the (large) separation and the lacking correction for recrossings, this constitutes a shooting algorithm in the vein of the one presented in Ref. 5.

"Shooting from the top" should prove particularly useful in cases where the cost of generating transition paths is high (as, e.g., in protein folding[10]) yet reasonable order parameters are known.[12–14] Arguably, this applies to quite a number of practically relevant cases, as attested by the success of algorithms such as forward-flux sampling,[6] transition interface sampling,[7] and milestoning,[11] whose efficiencies also rely on having reasonable foliations of the space between reactants and products.

SUPPLEMENTARY MATERIAL

See supplementary material for Figs. S1 and S2 which show the results of transition path analyses with misplaced shooting ranges for Langevin dynamics with barrier $B = 3\,k_BT$ and for overdamped Langevin dynamics with barrier $B = 10\,k_BT$, respectively.

ACKNOWLEDGMENTS

This work was supported by the Max Planck Society. K.O. is supported by the Building of Consortia for the Development of Human Resources in Science and Technology, MEXT, Japan.

[1] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, Annu. Rev. Phys. Chem. **53**, 291 (2002).
[2] P. R. ten Wolde, M. J. Ruiz-Montero, and D. Frenkel, J. Chem. Phys. **104**, 9932 (1996).
[3] C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, J. Chem. Phys. **108**, 1964 (1998).
[4] C. Dellago, P. G. Bolhuis, and D. Chandler, J. Chem. Phys. **108**, 9236 (1998).
[5] G. Hummer, J. Chem. Phys. **120**, 516 (2004).
[6] R. J. Allen, P. B. Warren, and P. R. ten Wolde, Phys. Rev. Lett. **94**, 018104 (2005).
[7] T. S. van Erp, D. Moroni, and P. G. Bolhuis, J. Chem. Phys. **118**, 7762 (2003).
[8] B. Peters and B. L. Trout, J. Chem. Phys. **125**, 054108 (2006).
[9] Z. F. Brotzakis and P. G. Bolhuis, J. Chem. Phys. **145**, 164112 (2016).
[10] J. Juraszek and P. G. Bolhuis, Biophys. J. **95**, 4246 (2008).
[11] A. K. Faradjian and R. Elber, J. Chem. Phys. **120**, 10880 (2004).
[12] K. Lindorff-Larsen, S. Piana, R. O. Dror, and D. E. Shaw, Science **334**, 517 (2011).
[13] R. B. Best, G. Hummer, and W. A. Eaton, Proc. Natl. Acad. Sci. U. S. A. **110**, 17874 (2013).
[14] R. B. Best and G. Hummer, Proc. Natl. Acad. Sci. U. S. A. **113**, 3263 (2016).
[15] R. B. Best and G. Hummer, Proc. Natl. Acad. Sci. U. S. A. **102**, 6732 (2005).
[16] P. G. Bolhuis and C. Dellago, Rev. Comput. Chem. **27**, 111 (2011).

[17] G. Menzl, A. Singraber, and C. Dellago, Faraday Discuss. **195**, 345 (2016).
[18] W. E and E. Vanden-Eijnden, J. Stat. Phys. **123**, 503 (2006).
[19] S. V. Krivov and M. Karplus, Proc. Natl. Acad. Sci. U. S. A. **101**, 14766 (2004).
[20] M. J. Abraham, D. van der Spoel, E. Lindahl, B. Hess, and GROMACS Development Team, GROMACS User Manual version 5.0.4., 2014 www.gromacs.org.
[21] N. Goga, A. J. Rzepiela, A. H. de Vries, S. J. Marrink, and H. J. C. Berendsen, J. Chem. Theory Comput. **8**, 3637 (2012).
[22] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).
[23] A. Ma and A. R. Dinner, J. Phys. Chem. B **109**, 6769 (2005).
[24] Y. LeCun, Y. Bengio, and G. Hinton, Nature **521**, 436 (2015).
[25] C. H. Bennett, "Molecular dynamics and transition state theory: The simulation of infrequent events," in *Algorithms for Chemical Computations*, ACS Symposium Series Vol. 46 (American Chemical Society, 1977), Chap. 4, pp. 63–97.
[26] D. Chandler, J. Chem. Phys. **68**, 2959 (1978).
[27] J. Daru and A. Stirling, J. Chem. Theory Comput. **10**, 1121 (2014).
[28] S. H. Northrup, S. A. Allison, and J. A. McCammon, J. Chem. Phys. **80**, 1517 (1984).
[29] R. R. Gabdoulline and R. C. Wade, Biophys. J. **72**, 1917 (1997).
[30] G. Schreiber, G. Haran, and H. X. Zhou, Chem. Rev. **109**, 839 (2009).